DOCTOR OF PHILOSOPHY

The identification and characterisation of proteins interacting with SUMO-ubiquitin hybrid chains

Anderson, Oliver

Award date:
2015

The identification and characterisation of proteins interacting with SUMO-ubiquitin hybrid chains

A thesis submitted in the fulfillment of the requirements for the degree of Doctor of Philosophy

by

Oliver Fraser Anderson

College of Life Sciences, University of Dundee, United Kingdom

September 2015

Ph.D. Supervisor: Professor Ron Hay

Declaration

I hereby declare that the following thesis is based on results of investigations conducted by me, and that this thesis is of my own composition. Work other than my own is clearly indicated in the text by reference to the relevant researchers or their publications. All references cited here have been consulted by me. This thesis has not, in whole or part, been previously accepted for a higher degree.

Oliver Anderson

I certify that Oliver Anderson performed the work for which this thesis is a record. The conditions of the relevant Ordinance and Regulations have been fulfilled.

Professor Ronald T. Hay

Acknowledgement

Firstly, I would like to thank my supervisor Professor Ron Hay for allowing me to undertake the research for this thesis in his laboratory, and for all his help and guidance over the last 3 years. I would also like to thank all the members of the RTH lab that I have had the pleasure of working alongside. In particular, I would like to thank Ellis Jaffray for all his technical assistance, Michael Tatham for all the discussions and help with proteomics, Federico Pelisch for help with microscopy, and Anne Seifert and Yili Yin for all their great advice.

I would also like to thank the Biotechnology and Biological Sciences Research Council for generously providing the funding for this research. My thanks must also go out to the all the people of the College of Life Sciences, University of Dundee for making it such a great place to work and learn.

Finally, Thank you to all my friends and especially to my family Mum, Dad, Sarah, Megan, Bethany, Morgan and Michael for the endless support over the last 3 years, without which none of this would be possible.

Contents

Table of Contents

List of figures

Abbreviations

53BP1 – Tumour suppressor proteins p53-binding protein 1

6His – Six histidine residue tag

Ac – Acidic residue

ATM – Ataxia telangiectasia mutated

ATP – Adenosine triphosphate

BARD1 – BRCA1-associated RING domain 1

BRCA1 – Breast cancer susceptibility type 1

BRCC36 – Lys-63-specific deubiquitinase BRCC36

C-terminus – Carboxyl terminus

C5orf25/SIMC1 – SIM containing protein 1

cDNA – Complementary DNA

DAPI – 4',6-diamidino-2-phenylindole

DDR – DNA damage response

Desi – DeSUMOylating isopeptidases

DMEM – Dulbecco's modified eagle medium

DNA – Deoxyribonucleic acid

DSBs – DNA double strand break

DSTT – Division of signal transduction

DTT – dithiothreitol

DUbs – Deubiquitinases

E. coli – *Escherichia coli*

E1 – Activating enzyme 1

E2 – Conjugating enzyme 2

E3 – Ligating enzyme 3

ECL – Enhanced chemiluminescence

FAT10 – HLA-F-adjacent transcript 10

FCS – Fetal calf serum

GFP – Green fluorescence protein

GST – Glutathione S-transferase

H2Ax – Histone H2Ax

HECT – Homologues to E6-AP C-terminus

HR – Homologous recombination

HRP – Horse radish peroxidase

Ig – immunoglobulin

IP – Immunoprecipitation

IPTG – Isopropyl β-D-1-thiogalactopyranoside

IR – Ionising radiation

IRIF – Ionising radiation-induced foci

JAMM – Jab1/Pab1/MPN domain containing proteases

kDa – Kilo Dalton

LB – Luria-Bertani medium

LDS – litium dodecyl sulphate sample loading buffer

LUBAC – Linear ubiquitin assembly complex

MBP – Maltose binding protein

MDC1 – Mediator of DNA damage checkpoint 1

Mdm2 – Mouse double minute 2 homologue

MIU – Motif interacting with ubiquitin

MJD – Machado-Joseph domain-containing proteases

MOPs – 3-morpholinopropane

MORC3 – MORC family CW-type zinc finger protein 3

MPN – JAMM/MPN type ubiquitin binding domain

MRN – MRE11-Rad50-NSB1 containing complex

N-terminus – amino terminus

NEDD8 – Neural precursor cell expressed, developmentally down-regulated 8

NEMO – NF-kappa-B essential modulator

NF-κB – Nuclear factor kappa-light-chain-enhancer of activated B cells

NHEJ – Non-homologous end joining

NHS – N-hyrdoxysuccinimide

Ni-NTA – Nitriloacetic acid

NLS – Nuclear localisation signal

No.T – No treatment

(P) – Phosphorylation

P53 – cell tumor antigen 53

PAGE – Polyacrylamide gel electrophoresis

PBS – phosphate buffered saline

PFA – paraformaldehyde

PIAS – Protein inhibitor of activated STAT

PLA – Proximity ligation assay

PML – Promyelocytic leukemia protein

PML-NBs – PML nuclear bodies

PPPDE – Permutated papin fold peptidase of double stranded RNA viruses and eukaryotes

Pre-SUMO – Precursor SUMO polypeptide

PTM – Post-translational modification

PVDF – Polyvinylidene

RanBP2 – RanGAP1 binding protein 2

RanGAP1 – GTPase-activating enzyme 1

RAP80/UIMC1 – Receptor associated protein 80/UIM containing protein1

RING – Really interesting new gene

RIPA – Radioimmunoprecipitation assay buffer

RNF111 – RING finger protein 111

RNF168 – RING finger protein 168

RNF4 – RING finger protein 4

RPA – Replication protein A

RPM – Revolutions per minute

RPS27A – Ubiquitin-40S ribosomal protein S27a precursor

S. cerevisiae – Saccharomyces cerevisiae

S2 – SUMO2

SAE1/SAE2 – SUMO activating E1 enzyme

SDS – Sodium dodecylsulphate

SENP – Sentrin/SUMO-specific protease

SILAC – Stable isotopically labelled amino acids in cell culture

SIMs – SUMO interaction motif

Siz – SAP and mix domains

Slx – Synthetic lethal X gene

SMEM – Minimum essential eagles medium spinner modification

SP100 – Nuclear antigen SP100

STUbL – SUMO targeted ubiquitin E3 Ligase

SUMO1/2/3/4 – Small ubiquitin-like modifier 1/2/3/4

TBS – Tris-buffered saline

TCA – Trichloroacetic acid

TCEP – Tris(2-carboxyethyl)phosphine

TEV – Tobacco etch virus

TOPORs – Topoisomerase I-binding RING protein

Tris – Tris(hydroxymethyl)aminomethane

Ub – Ubiquitin

Uba1 – Ubiquitin-like modifier-activating enzyme 1

UBA52 – Ubiquitin-60S ribosomal protein L40 precursor

Uba6 – Ubiquitin-like modifier-activating enzyme 6

UBB – Poly-ubiquitin B

UBC – Poly-ubiquitin C

Ubc – Ubiquitin conjugating domain

Ubc13 – Ubiquitin conjugating enzyme 13

Ubc9 – Ubiquitin conjugating enzyme 9

UBDs – Ubiquitin binding domains

Ube2v2 – Ubiquitin conjugating enzyme E2 variant 2

Ubls – Ubiquitin-like modifiers

Uev – E2-variant enzyme

UHC – Ubiquitin C-terminal hyrdolases

UIM – Ubiquitin interaction motif

USP – Ubiquitin-specific proteases

USP2 – Ubiquitin-specific protease 2

USP2 CaD – Active catalytic domain of USP2 (residues Asn258-Met605)

USP25 – Ubiquitin-specific protease 25

USP28 – Ubiquitin-specific protease 28

USPL1 – Ubiquitin-specific protease like 1

yH2Ax – Phosphorylated histone H2Ax

Znf – Zinc finger

β-GF - β-grasp fold

Ψ – Hydrophobic residue

Abstract

A wide variety of cellular processes are regulated via the post-translation modification of substrate proteins by either ubiquitin or the small ubiquitin-like modifier, SUMO. In recent years, points of convergence between the once thought distinct and independent SUMO- and ubiquitin-conjugation pathways have been identified. The ubiquitin E3 ligase RNF4 catalyses the addition of ubiquitin modifications to previously SUMO modified proteins. This results in hybrid chains of SUMO and ubiquitin decorating target substrates. Do these SUMO-ubiquitin hybrid chains act as unique signals, distinct from the SUMO and ubiquitin chains that make them? Are there proteins that recognise them as such? The first part of this thesis details a protocol to identify and characterise the affinity of cellular proteins for 'baits' consisting of hybrid chains of SUMO and ubiquitin. This affinity chromatography based approach when coupled with high resolution mass spectrometry identified 30 proteins from HEK 293 cellular extracts that were putatively identified as showing affinity specifically for SUMO-ubiquitin hybrid chains. Validation of this approach comes from the identification of the recently postulated SUMO-ubiquitin hybrid chain interacting protein RAP80 amongst the proteins identified as putative SUMO-ubiquitin hybrid chain interacting proteins. SUMO-ubiquitin hybrid chains are then evaluated in an *in vivo* context. A proximity ligation assay was developed to probe the association between SUMO, ubiquitin and RAP80, suggesting a tight association between the three after DNA damage inducing stimuli. Interesting, the SUMO targeted ubiquitin E3 ligase RNF4 was shown not to be required for the recruitment of RAP80 to sites of DNA damage. The association between RAP80, SUMO, and ubiquitin was probed further by immunoprecipitaion of RAP80 from cell extracts after a SUMO specific proteases treatment. This SUMO specific protease treatment resulted in the loss of high molecular weight ubiquitin conjugates from the RAP80 associated material suggesting that some RAP80 associated material is carrying SUMOylation dependent ubiquitin modifications. A SILAC based gel shift assay utilising both ubiquitin and SUMO-specific proteases was then developed to identify proteins that may anchor SUMO-ubiquitin hybrid chains.

Although technically challenging, S100A8 was identified as a protein that after the action of ubiquitin and SUMO proteases shifts in a pattern consistent with that expected of a protein that may anchor a SUMO-ubiquitin hybrid chain.

1 Introduction

Protein function is not only controlled by amino-acid sequence, but by a multitude of post-translational modifications (PTMs) that can alter the function of the modified protein. PTMs come in many forms ranging from proteolytic processing, to the addition of molecular adducts to specific amino acids. Common PTMs of the latter class include phosphorylation, methylation, acetylation and glycosylation. However, a growing number of PTMs are themselves polypeptides that provide scope for a very complex, interactive and often reversible system of protein regulation. Protein ubiquitylation, or the covalent modification of a protein by ubiquitin, is the classical example of this polypeptide class PTM.

1.1 Ubiquitin

1.1.1 Ubiquitin and the ubiquitin-like modifiers

Ubiquitin is a small 76 amino acid 8.5 kDa protein that is ubiquitously expressed and highly conserved in eukaryotes, differing by only a few amino acids from yeast to humans (Goldstein et al., 1975). Four human genes encode ubiquitin; UBB, UBC, UBA52, and RPS27A (Redman & Rechsteiner 1989). UBA52 and RPS27A express single ubiquitin moieties fused to ribosomal polypeptides, whereas UBB and UBC express linear fusion proteins containing multiple ubiquitin moieties (Finley et al., 1989). Discovered in the late 1970's as a trigger for non-lysosomal protein degradation (Ciechanover et al 1978), it is now known that the ubiquitylation machinery recognises and targets thousands of proteins, with implications on almost all cellular systems; the most well characterised action of protein ubiquitylation being the marking of proteins for degradation by the 26S proteasome (Hershko et al., 1984).

Ubiquitin belongs to a wider family known as the ubiquitin-like modifiers (Ubls). Since the identification of ubiquitin in the 1970's the mammalian Ubl family of proteins has grown to contain around 20 members including NEDD8, SUMO1, SUMO2, SUMO3, and FAT10 (Kerscher et al., 2006). Members of the Ubl family of proteins share variable levels of sequence similarity to

ubiquitin, but all contain a conserved β-grasp fold (β-GF) that is characteristic of the founding member ubiquitin (Hochstrasser 2009).

1.1.2 The ubiquitin conjugation pathway and associated enzymes

Protein ubiquitylation involves an enzymatic cascade resulting in the production of an isopeptide bond between the C-terminal Gly76 of ubiquitin and most commonly the ε-amino group of a lysine residue contained in a target substrate (**Figure 1.1-1**). The enzymatic cascade typically involves three distinct sets of enzymes; firstly, in an ATP-dependent step, a thioester bond is formed between the C-terminal Gly76 of ubiquitin and a cysteine of an E1 activating enzyme. Subsequently, ubiquitin is transferred to the active site cysteine of an E2-conjugating enzyme before finally being transferred to a lysine residue located in a target protein via the action of an ubiquitin E3 ligase.

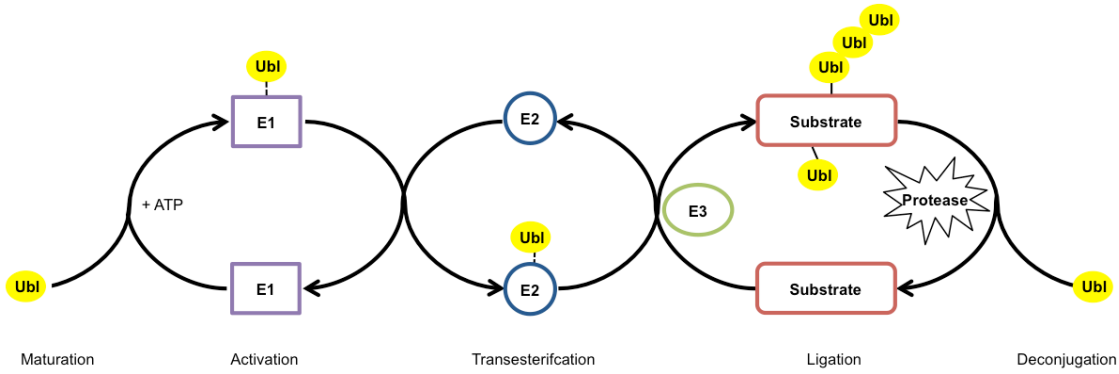


Figure 1.1-1 The Ubl conjugation cycle.

A 3-step enzymatic cascade involving E1, E2 & E3 enzymes resulting in the conjugation of Ubl family members to range of target substrates, modification of Ubl family members results in the formation of polymeric chains of some Ubl proteins. Proteases antagonise E3 enzymes by facilitating the removal and recycling of the Ubls from target substrates.

The human genome encodes two ubiquitin E1 activating enzymes; Uba1 and the more recently identified Uba6 (Jin et al., 2007). Uba1 is thought to be the apical E1 enzyme for the majority of ubiquitylation event *in vivo,* as such, it has been shown that Uba1, but not Uba6, is critical for the formation

of ionising radiation induced foci formation at the sites of DNA damage (Moudry et al., 2012).

Around 40 ubiquitin E2 enzymes have so far been identified. E2 enzymes are recognisable by a highly conserved ~150 amino acid ubiquitin-conjugating domain (Ubc), containing the active-site cysteine to which ubiquitin becomes conjugated. Along with the conserved Ubc domain, most E2 enzymes also contain N- and/or C- terminal extension (Burroughs et al., 2008). E2 enzymes must perform a number of functions during ubiquitin conjugation. Firstly, they must act to recognise and then transiently receive ubiquitin from an E1 enzyme; this results in the formation of an E2-ubiquitin thioester. Secondly, this E2-ubiquitin thioester must then interact with an E3 ligase enzyme to catalyse the ubiquitylation of a target substrate (**Figure 1.1-1)**. Additional E2-variant (Uev) enzymes have been identified which also contain the Ubc domain (Burroughs et al., 2008). These enzymes lack the catalytic active site cysteine required for ubiquitin conjugation and so are thought to be conjugation dead pseudo enzymes. However some E2-variants are known to aid E2 enzymes in positioning ubiquitin for linkage-specific conjugation. For example, the E2/E2-variant pair Ubc13/Ube2v2, Ube2v2 has been observed to be responsible for recognising and binding the acceptor ubiquitin and positioning the acceptor K63 lysine for attack on the linkage between Ubc13 and the donor ubiquitin which is held in place by the E3 ligase RING finger protein 4 (RNF4) (Branigan et al., 2015).

Approximately 600 ubiquitin E3 enzymes have been identified in humans. There are two main classes of ubiquitin E3 enzyme: RING (Really Interesting New Gene) domain containing proteins, and HECT (Homologues to E6-AP C-terminus) domain containing proteins (Metzger et al., 2012). RING domain containing ubiquitin E3 ligases function to catalyse the direct transfer of ubiquitin from an E2-ubiquitin thioester to a target protein. This mechanism was first observed in crystal structures of the E2-ubiquitin thioester, UbcH5a-Ub with the RING domain of RNF4 (Plechanovova et al., 2011). In contrast, HECT domain containing ligases contain an active-site

cysteine residue that accepts ubiquitin from an E2 prior to substrate modification (Hung et al., 1999).

1.1.3 Ubiquitin chain types

The ubiquitylation of target substrates can result in a diverse set of signals (**Figure 1.1-2**). Proteins can be singly mono-ubiquitiylated, multiply mono-ubiquitylated, or due to the presence of seven lysine residues (K6, K11, K27, K29, K33, K48, K63) (**Figure 1.1-3, A**) contained within the sequence of ubiquitin, ubiquitin can modify other ubiquitin proteins resulting in the formation of polymeric chains of ubiquitin (poly-ubiquitin) (Kultha & Kommander 2012) (**Figure 1.1-2, B**). Furthermore, linear poly-ubiquitin can also be formed through the N-terminal amino group of Met1 (Iwai et al., 2009) (**Figure 1.1-2, B**). All eight ubiquitin linkage types have been observed *in vivo* (Peng et al., 2003; Xu et al., 2009). It is thought that different combinations of ubiquitin E2/E3 pairs are responsible of the addition of specifically linked poly-ubiquitin chains to specific substrate proteins. K48 linked and K63 linked poly-ubiquitin chains are the most well characterised examples of poly-ubiquitin. K48 linked poly-ubiquitin chains are thought to make up around 30% of the total number of polymeric ubiquitin chains present in any cell at any one time (Xu et al., 2009). The addition of a K48 linked poly-ubiquitin chain, of at least 4 ubiquitin moieties, to any protein functions to target that protein for destruction by the 26S proteasome (Thrower et al., 2000). K63 linked poly-ubiquitin chains have been observed to mark proteins from a wide variety of cellular processes including the DNA damage response, but do not appear to have a function in targeting proteins for destruction by the proteasome, instead this linkage type is thought to act as a wider signalling mechanism (Hofmann & Pickart 1999). In comparison to K48 linked and K63 linked poly-ubiquitin chains, little is known about the remaining linkage types. K11 linked poly-ubiquitin chains are less well studied but under certain circumstances are almost as abundant as K48 linked poly-ubiquitin chains (Xu et al., 2009; Matsumoto et al., 2010), and have also been observed to have a role in proteasomal degradation (Wickliffe et al., 2011). K6 linked poly-ubiquitin chains can be catalysed by the breast cancer susceptibility type 1

(BRCA1) and BRCA1-associated RING domain protein 1 (BARD1) (BRCA1-BARD1) E3 ligase complex and so may have a role in the DNA damage response (Wu-Baer et al., 2003). Several putative roles have been ascribed to K27, K29, and K33 linked poly-ubiquitin chains, with all three linkage types appearing to be important signalling mechanisms in T-cells (Chastagner et al., 2006; Peng et al., 2011). Interestingly, two of the four genes that encode ubiquitin in the human genome are essentially Met1 linked poly-ubiquitin chains and as such act as precursors for all poly-ubiquitin linkage types in the cell. However, a large 600 kDa protein complex known as the Linear Ubiquitin Assembly Complex (LUBAC) has also been observed to catalyse the addition of Met1 linked poly-ubiquitin chains to proteins including NF-kappa-B essential modulator (NEMO) and other substrates important for the activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) which plays a key role in the cellular immune response (Iwai et al., 2009; Tokunaga et al., 2009; Niu et al., 2011).

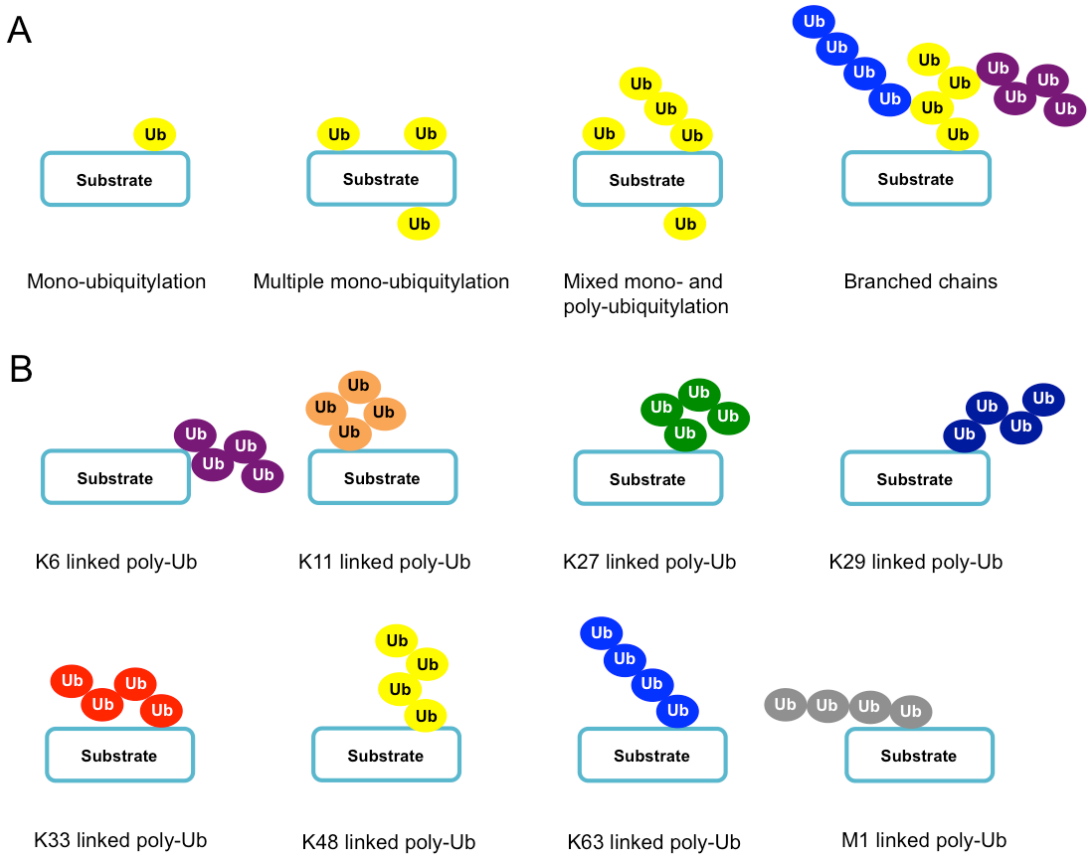


Figure 1.1-2 Ubiquitin modifications.

A) Depending on the number and availability of lysine residues, substrate proteins can be mono-ubiquitylated, multiply mono-ubiquitylated, poly-ubiquitylated or combinations of both. Substrate proteins can also anchor heterologous poly-ubiquitin modifications known as branched chain modifications.
B) Branched chain modifications can be made up from any combination of the eight different poly-ubiquitylation types. K6 linked pUb is denoted by purple Ub moieties. K11 linked pUb is denoted by orange Ub moieties. K27 linked pUb is denoted by green Ub moieties. K29 linked pUb is denoted by dark blue Ub moieties. K33 linked pUb is denoted by red Ub moieties. K48 linked pUb is denoted by yellow Ub moieties. K63 linked pUb is denoted by light blue Ub moieties. M1 linked pUb is denoted by grey Ub moieties.

Differences in topology of the eight different ubiquitin linkage types allow for a wide range of signals to be generated by poly-ubiquitin chains. Homotypic poly-ubiquitin chains are thought to primarily adopt one of two structural conformations; a compact confirmation where the ubiquitin moieties composing the chain interact with each other via hydrophobic surface on ubiquitin along with the linkage site; or an open confirmation where the ubiquitin moieties comprising the chains are only interfacing via the linkage site (Komander & Rape 2012). It has been reported that K48 linked ubiquitin moieties form a compact conformation (Tenno et al., 2004). The compact nature of linkages suggests that some of the binding surface of ubiquitin may be closed off to interacting partners (Dikic et al., 2011). Conversely, K63 and Met1 linked ubiquitin moieties form open conformations (Tenno et al., 2004; Varadan et al., 2004). The open conformation allows these linkage types to be more flexible with a high degree of conformational freedom that is likely exploited by any interacting partners of these chain types (Tenno et al., 2004; Varadan et al., 2004).

1.1.4 Recognition of ubiquitin chains types by ubiquitin binding domains

Ubiquitin contains several distinctive interaction surfaces (**Figure 1.1-3, B**). Ubiquitin is predominantly recognised via a hydrophobic patch centred on Ile44, known as the Ile44 patch. It is predominantly through interactions with this hydrophobic patch that K48 linked poly-ubiquitin chains form a compact

conformation. However, hydrophobic patches including the Ile36 patch, Phe4 patch and the TEK box are also known to act as sites for non-covalent interactions with ubiquitin (Dikic et al., 2011) (**Figure 1.1-3, B**). Indeed it has been observed that the closed conformation of K48 linked poly-ubiquitin chains can also be facilitated via the Ile36 patch, freeing up the Ile44 patch to non-covalent interactions with other ubiquitin interacting partners (Eddins et al., 2007).

In recent years a wide variety of proteins have been found to contain distinct domains that facilitate a non-covalent interactions with ubiquitin, known as ubiquitin-binding domains (UBDs). The UBD family is composed of a wide variety of domains that between them have the ability to interact with all types of ubiquitin modifications from mono-ubiquitin through all the specific poly-ubiquitin chain types. One of the best characterised types of UBD is the ubiquitin interaction motif (UIM). UIMs were first identified as the UBD contained within the S5a (RPN10) subunit of the 26S proteasome (Young et al., 1998; Hofmann & Flaquet., 2001). The conserved UIM motif is composed of 15 residues e-e-e-e-x-Ψ-x-x-A-x-x-x-S-x-x-e (where "e" is any negatively charged residue and Ψ is a hydrophobic residue) and have been observed to form a α-helice (Hofmann & Flaquet., 2001;Swanson et al., 2003). UIMs are some times referred to as the "LALAL" motif due to its Leu and Ala rich core region, while a block of four preferentially acidic residues precedes this core region (Hofmann & Flaquet., 2001; Swanson et al., 2003). Single UIMs show low affinity for mono-ubiquitin. However, many proteins identified containing UIMs have been shown to contain multiple UIMs, and show high-affinity for poly-ubiquitin, including the BRCA 1A complex component, RAP80 that contains two UIM domains. These two UIMs allow RAP80 to interact non-covalently with a dimer of K63 linked ubiquitin (**Figure 1.1-3, C**). UIMs are thought to interact with Ile44 patch of ubiquitin (Swanson et al., 2003). Many of the UBDs identified so far are known to interact with this hydrophobic patch of ubiquitin including the 26S proteasome (Dikic et al., 2011). However, there are exceptions, including the A20 type Znf domain of RAB5. The A20 type Znf domain of RBA5 forms interactions not only with the Ile44-containing surface, but also the TEK box and a polar surface of ubiquitin centred on

Asp58 (Lee et al., 2006) (**Figure 1.1-3, B**). The list of UBDs is likely to be incomplete, with new domains being regularly described.

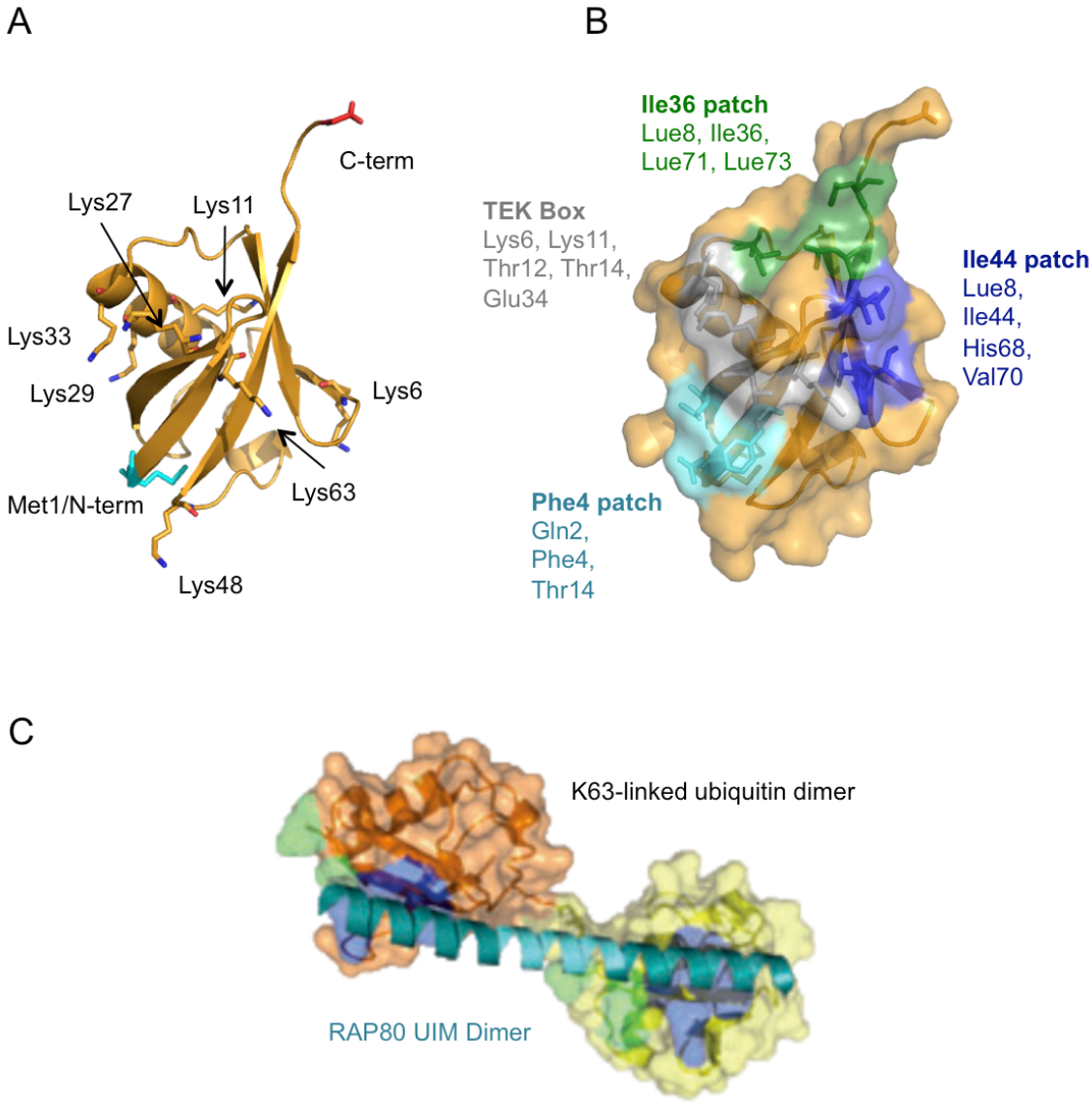


Figure 1.1-3 Interacting with ubiquitin.

A) Structure of ubiquitin highlighting the seven internal lysine and Met1 residues involved in covalent poly-ubiquitin modifications. Met1/N-terminus (N-term) highlighted in cyan. C-terminus (c-term) is highlighted in red. Protein Data Bank identifier 1ubq.
B) Structure of ubiquitin highlighting hydrophobic non-covalent interaction surfaces. Ile36-patch highlighted in green. Ile44-patch highlighted in blue. TEK-box highlighted in grey. Phe4-patch highlighted in cyan.
C) Structure of the non-covalent interaction between a K63-linked ubiquitin dimer (orange and yellow) and a RAP80 UIM dimer (cyan) via the Ile44-patch (Blue) located on each of the ubiquitin moieties (Komander & Rape 2012). Protein Data Bank identifier 3a1q.

1.1.5 Ubiquitin deconjugating enzymes

Ubiquitylation is a dynamic, and importantly, a reversible process. The removal of mono- and poly-ubiquitin is facilitated by a large group of enzymes known as deubiquitinases (DUbs) (Nijman et al., 2005). DUbs carry out a number of essential roles including processing ubiquitin precursors, antagonising ubiquitin E3s (**Figure 1.1-1**) and recycling K48 linked poly-ubiquitin chains from proteins targeted to the proteasome just prior to the target protein being destroyed (Kommander et al., 2009). The human genome encodes approximately 100 DUb enzymes, which can be classified into five subfamilies; Ubiquitin C-terminal hydrolases (UHC), Ubiquitin-specific proteases (USPs), Otubain domain ubiquitin binding proteins (OTU), Machado-Joseph domain-containing proteases (MJD), and the Jab1/Pab1/MPN domain containing proteases (JAMM). Four of these subfamilies UHC, USP, OTU, and MJDs are classed as cysteine-specific proteases with the JAMM family classed as Zinc metalloproteases (Nilman et al., 2005). The USP family is the largest single subfamily of DUbs, containing 56 members, thus the majority of reported DUb enzymes are cysteine-specific proteases. Cysteine specific proteases contain a catalytic triad of Cys-His-Asp/Asn. USP2 is a member of the USP family of DUb enzymes and is highly active both *in vitro* and *in vivo* (Renatus et al., 2006). USP2 is thought to play a role in regulating p53 “the guardian of the genome” by deubiquitylating the p53 regulating protein Mdm2 (Stevenson et al., 2007). Mdm2 is a ubiquitin E3 ligase important for the proteasome mediated degradation of p53 (Brady et al., 2005). Deubiquitylaion of Mdm2 by USP2, and USP7 another member of the USP family of DUbs, leads to the stabilisation and activation of Mdm2 and thus inhibition of p53 (Stevenson et al., 2007). USP2, consisting of 605 amino acids, is a relatively small USP protein, as many of the USP family proteases are composed of ~1000 amino acids. The catalytic triad of USP2, Cys276-His557-Asp574, is located in its C-terminal USP domain, which is a characteristic of the USP family of proteases (Renatus et al., 2006). The JAMM domain containing proteases are classed as Zn metalloproteases and conform to a catalytic active-site consensus motif of Glu-X-[N]-His-His-X(10)-Asp. The BRCA 1A complex component BRCC36, a JAMM domain

containing protease, is thought to have a role in deconjugating K63 linked poly-ubiquitin from substrate proteins at the sites of DNA damage (Shao et al., 2009A). DUbs often have discrete domains that allow them to recognise and interact non-covalently with ubiquitin.

1.2 SUMO

1.2.1 SUMO the small ubiquitin-like modifier

In the years following the identification of ubiquitin several other ubiquitin like proteins were identified including the small ubiquitin-like modifier (SUMO). SUMO was identified in the mid 1990's as a polypeptide adduct of the GTPase-activating enzyme RanGAP1, regulating the transport of RanGAP1 from the cytoplasm to the nuclear pore complex (Matunis et al., 1996; Mahajan et al., 1997). Since its identification, SUMO has been shown to alter the function of a wide variety of proteins in several essential cellular processes such as cellular localisation, transcription, heat shock response, and the DNA damage response (Gill 2003; Golebiowski et al., 2009; Galanty et al., 2009).

The SUMO family of proteins share limited sequence identity with ubiquitin (~20%) but contain a similar β-GF structure (Hochstrasser 2009). In humans, there are 3 active forms of SUMO; SUMO1, SUMO2 and SUMO3. SUMO2 and SUMO3 are nearly identical, only differing by three amino acids in their N-terminal region and have not been functionally differentiated; as such they form the SUMO2/3 subfamily (referred to as SUMO2 from now on) sharing 46-48% sequence identity with SUMO1 (Saitoh & Hinchey, 2000; Tatham et al., 2001). Paralog-specific roles have been ascribed to both SUMO1 and SUMO2, with some substrates preferentially modified by either SUMO1 or SUMO2; RanGAP1, for example is preferentially modified by SUMO1 (Saitoh & Hinchey, 2000). However, SUMO1 and SUMO2 share specificity for a large number of substrates and a degree of functional redundancy exists between them (Vertegaal et al., 2006). For example, mice lacking SUMO1 are still viable suggesting that SUMO2 can functionally compensate for SUMO1 in some instances (Zang et al., 2008). Interestingly a more recent study has suggested that SUMO3 null mice are viable but SUMO2 null mice were embryonic lethal (Wang et al., 2014). The authors of this report suggested that this difference between SUMO2 and SUMO3 was not due to a functional difference between the two proteins, but rather because SUMO2 is the predominantly expressed SUMO isoform during embryogenesis (Wang et al.,

2014). The majority of SUMO1 in cells is found conjugated to proteins under normal growth conditions, but interestingly, free pools of SUMO2 exist under normal cell growth conditions. Under cellular stress condition such as heat shock, rapid conjugation of this pool of free SUMO2 to a variety of proteins is required for cellular survival (Golebiowski et al 2009). Humans contain a fourth SUMO gene, SUMO4, the product of which shares 85% sequence identity with SUMO2. SUMO4 has been suggested to have a role in signalling stress response in kidney cells and a role in the pathogenesis of type-1 diabetes (Brohen et al., 2004). However, it is unclear if this protein can be actively conjugated to proteins and so it is thought SUMO4 may in fact be a pseudo gene (Owerbach et al., 2005).

1.2.2 SUMO conjugation pathway

SUMO1 and SUMO2, like ubiquitin, are conjugated to target substrates through an enzymatic cascade involving a distinct set of E1, E2 and E3 enzymes characteristic of the Ubl family of proteins (**Figure 1.1-1**). All SUMO proteins are expressed as inactive precursor proteins before SUMO-specific proteases act to cleave the C-terminal regions of SUMO1, SUMO2 and SUMO3 to reveal a Gly-Gly motif capable of covalent linkage to the ε-amino side chain of a lysine residue found within an acceptor substrate (Hay 2005). The presence of Pro90 in SUMO4 impairs the cleavage of its C-terminal, thus inhibiting the activation and conjugation of SUMO4 to target substrates (Owerbach et al., 2005). A heterodimer of SAE1 and SAE2 forms the SUMO activating E1 enzyme. The ATP-dependent action of SAE1/SAE2 results in the formation of a thioester bond between a cysteine residue of SAE2 and the C-terminus of SUMO. SUMO is then transesterified to the SUMO conjugating E2 enzyme. Unlike the ubiquitin conjugation system where ~40 E2 enzymes have been identified, only one functional E2 enzyme has been identified in the SUMO conjugation pathway, Ubc9. Another differentiating factor in the SUMO conjugation pathway is that unlike the ubiquitin conjugation pathway, Ubc9 is able to recognise and transfer SUMO directly to target proteins in the absence of an E3 ligase enzyme (Desterro et al., 1997). This is due to Ubc9's ability to recognise lysine residues capable of accepting SUMO modification. Ubc9 is so functionally important for SUMO conjugation that mice deficient in

Ubc9 are embryonic lethal, highlighting the functional importance of the SUMO system to cell survival (Nacerddine et al., 2005). Although SAE1/SAE2 and Ubc9 are sufficient to form an iso-peptide bond between the C-terminus of SUMO and the ε-amino group of an acceptor lysine the reaction is very inefficient. Like the ubiquitin conjugation pathway, SUMO E3 ligase enzymes act to increase the rate of conjugation. Less than 20 SUMO E3's have been identified compared to the approximately 600 ubiquitin E3's. However, more than 1600 proteins are known to be SUMO substrates and more than 4000 SUMOylation sites have been identified (Tammsalu et al., 2014; Hendricks et al., 2014). This suggests that a relatively small number of SUMO E3 ligases are capable of modifying large numbers of distinct SUMO substrates. The first group of SUMO E3s identified were the Protein inhibitor of activated STAT (PIAS) proteins (Wong et al., 2004). The PIAS proteins are functionally similar to the RING-domain containing ubiquitin E3 ligases. Four human PIAS genes have been identified to encode SUMO E3 ligases; PIAS1, PIAS2 (PIASx), PIAS3, and PIAS4 (PIASy). The PIAS proteins have been shown to SUMO modify a wide variety of proteins involved in a number of different cellular processes. Interestingly, however, the substrate specificity of the PIAS proteins has been shown to be low, thus it is unclear how exactly these proteins actively target substrates for SUMOylation. Unlike the PIAS E3 ligase, some SUMO E3s have no comparable ubiquitin counterparts. RanBP2 was among the first SUMO E3 ligase enzymes to be identified, but is neither comparable to RING or HECT type ubiquitin E3s (Pichler et al., 2004). RanBP2 has both SUMO1 and SUMO2 E3 ligase activity, modifying SP100 with SUMO1 but preferentially modifying PML with SUMO2 *in vivo* (Lallemand-Breitenbach et al., 2001; Tatham et al., 2004). This apparent SUMO paralog-specific activity is due to the way in which RanBP2 recognises and interacts with the Ubc9-SUMO1 thioester when compared to the Ubc9-SUMO2 thioester. The E3 ligase activity of RanBP2 is contained within a 100 amino acid region that can be broken down into three distinct motifs IR1, M, and IR2. IR1 has been observed to interact strongly with Ubc9 whereas M-IR2 has been observed to interact with SUMO1. Interestingly RanBP2 does not interact with SUMO2 alone. Thus, the Ubc9-SUMO1 thioester is bound strongly by tandem interaction of the IR1, M, and IR2 containing domains

interacting with both the Ubc9 and SUMO1 moieties, whereas Ubc9-SUMO2 thioester is bound more weakly solely via the interaction between the IR1 domain and Ubc9 (Tatham et al., 2004). Thus the relative affinities for Ubc9-SUMO1 and Ubc9-SUMO2 thioesters are thought to be responsible in part for the paralog specificity shown by RanBP2. The region of RanBP2 that binds Ubc9 is on the opposite binding surface to the cysteine that forms a thioester with SUMO. It is thought that this allows RanBP2 to position the Ubc9-SUMO thioester for nucleophilic attack (Reverter and Lima 2005).

1.2.3 Consensus SUMO modification motif

Initially, SUMO conjugation was thought to be primarily targeted to lysine residues contained within a consensus SUMO modification motif, Ψ-K-x-E or Ψ-K-x-D (where Ψ is any bulky hydrophobic residue and x is any residue) (Rodriguez et al., 2001). Many SUMO modification sites were thereafter identified in proteins containing such consensus motifs. Later, an inverted SUMO consensus motif, E-x-K-Ψ or D-x-K-Ψ, was identified by a small-scale proteomic study (Matic et al., 2008). Advances in proteomics have greatly increased our understanding of SUMO conjugation sites. Large-scale proteomic analysis of SUMO sites has led to a better understanding of potential consensus SUMO modification motifs. A recent study identified 1002 distinct SUMO2 modified sites of which 70% conformed to either the forward or inverted consensus SUMO modification motif (Tammsalu et al., 2014). 90% of the forward containing consensus motifs were observed to contain a Glu at the +2 position suggesting that Ψ-K-x-E is the preferred variant of the forward consensus motif. Glu and Asp were shown to be equally likely to be present in -2 position of the inverted consensus motif, but, surprisingly, hydrophobic residues did not take preference in the +1 position. Interestingly, a fusion of forward and inverted motifs was also observed containing Glu or Asp in both the -2 and +2 positions (Tammsalu et al., 2014).

Polymeric modification of Ubls is not restricted to ubiquitin and its eight separate internal modification sites. An internal forward consensus SUMO conjugation motif was identified at the N-terminally located K11 (V10-**K11**-T12-E13) of SUMO2 allowing the formation of poly-SUMO2 chains and

increasing the complexity of signalling via the SUMO modification system (Tatham et al., 2001). SUMO1 does not contain this motif and is not thought to readily form poly-SUMO1 chains *in vivo,* however, SUMO1 has been suggested to "cap" the end of poly-SUMO2 chains, acting to halt chain polymeric SUMO chain formation (Tatham et al., 2001). Although K11 of SUMO2 is the primary acceptor lysine for modification by SUMO, modifications of K5 and K7 have also been reported (Tammsalu et al., 2014). This suggests a more complex branched SUMO modification landscape may exist. As of yet no biological role has been suggested for these atypical SUMO2 polymers (Tammsalu et al., 2014).

1.2.4 SUMO interaction motifs

The SUMOylation of proteins can be recognised non-covalently by other proteins containing specialised non-covalent binding motifs known as, SUMO interaction motifs (SIMs). In this way SUMOylation can aid in protein-protein interactions. This is observed most strikingly in PML nuclear bodies (PML-NBs) where the non-covalent interactions between SUMO modified PML and associated proteins acts to nucleate the formation of PML-NBs. If the SIMs of PML are mutated out the SUMOylation of PML fails to nucleate PML-NBs (Shen et al., 2006A). From structural studies it is known that SIMs form β-strands, which in turn form non-covalent hydrophobic interactions with a structural groove conserved in all three active human SUMO isoforms. This groove is situated between the second β-strand and the first alpha helix of SUMO (Song et al., 2005). All SIM domains are thought to interact with this same region of SUMO. The most well characterised "classical" SIM type motifs conform to the consensus sequence V/I/L-V/I/L-x-V/I/L or V/I/L-x-V/I/L-V/I/L, with groups of acidic residues or phosphorylated serine residues often observed in close proximity to the hydrophobic core SIM motif. These acidic or phosphorylated (negatively charged) residues are thought to help stabilise the interaction between the SIM motif and SUMO, by interacting with a positively charged lysine residue located proximal to the second β-strand and the first alpha helix of SUMO (Hecker et al., 2006). The location of these negatively charged amino acids is thought to affect the orientation and the affinity of the SIM-SUMO interaction, by allowing SIMs to bind SUMO in a

parallel or anti-parallel orientation with respect to the second β-strand of SUMO (Reverter & Lima 2005; Song et al., 2005). Interestingly, a "high fidelity" SIM type domain has been described conforming to the consensus sequence V/I/L/F/Y-V/I-D-L-T, with V-I-D-L-T representing the idealistic high fidelity SIM (Sun and Hunter 2012). The presence of Val, Ile, and Thr, which can more easily form β-strands gives this high fidelity SIM type a very rigid structure well suited to interacting with all SUMO isoforms (Sun and Hunter 2012).

Non-canonical SIM domains have also been described in the literature. Recent evidence has suggested the MYM-type zinc finger containing proteins ZMYM2 and ZMYM3 contain a potential new class of SIM domain (Guzzo et al., 2014). The authors of the study report that a single MYM-type zinc finger can interact with SUMO2 in a 1:1 stoichiometry, through the same surface as the classical type SIM domain. Five human proteins have been described to contain MYM-type zinc finger; ZMYM2, ZMYM3, ZMYM4, ZMYM5 and ZMYM6. The MYM-type zinc fingers in these proteins confirm to the consensus sequence 'Cys-X_2-Cys-X_{19-22}-Cys-X_3-Cys-X_{13-19}-Cys-X_2-Cys-X_{19-25}-Cys-X_3-Cys' (Guzzo et al., 2014). MYM-type zinc fingers were reported to aid the recruitment of ZMYM2 to SUMOylated PML in PML-NBs (Guzzo et al., 2014). ZMYM3 has been reported to show affinity for SUMO2, however, it is still unclear whether the affinity for SUMO2 shown by ZMYM3 is due to the presence of MYM-type zinc fingers (Guzzo et al., 2014). It is interesting to note that ZMYM2, ZMYM3, ZMYM4 and ZMYM6 all contain multiple MYM-type zinc fingers. It will be interesting to see in the future whether this allows these proteins to interact with polymeric SUMO chains. Other non-canonical SIM like domains have been described. For example, the ubiquitin RING type E3 ligase Herpes simplex virus type 1 early intermediate protein ICP0, is known to contain SIM like sequence (SLS). The SLSs of IPC0 are required for the IPC0 induced ubiquitin-dependent degradation of SUMOylated PML (Boutell et al., 2011).

The list of proteins known to contain SIMs grows rapidly as does our understanding of how these proteins interact with SUMO non-covalently. It is

interesting to note that, as in the ubiquitin system, the majority of SIM and SIM-like domains that have thus far been described all preferentially interact with the same binding surface of SUMO. How SUMO interacting proteins preferentially recognise SUMO modifications of one proteins from another is not clear. Certainly, in some cases it may involve secondary interactions with other proteins or the SUMO modified protein itself.

1.2.5 SUMO specific proteases

SUMOylation is a reversible modification. As with the DUb enzymes of the ubiquitin system, the SUMO system has a distinct set of enzymes that can actively remove SUMO from target substrates, known as deSUMOylation. In humans, nine SUMO isopeptidases have so far been identified, which can be classed into three families; the Sentrin/SUMO-specific proteases (SENPs), the DeSUMOylating isopeptidases (Desi), and the USPL1-related isopeptidases.

There are six active SENPs in the mammalian system; SENP1, SENP2, SENP3, SENP5, SENP6, and SENP7. All six active SENP enzymes contain a highly conserved catalytic cysteine-containing protease domain located in their C-terminal region. This region contains a Cys-His-Asp catalytic triad (SENP1; Cys-603, His-533, and Asp-550, Shen et al., 2006B) that is critical for the activity of the SENP enzymes. Hence, mutational analysis of these residues has been shown to disrupt the catalytic activity of the SENP enzymes (Shen et al., 2006B; Chung et al., 2010; Haindl et al 2008). The SENPs can be further characterised into 3 subfamilies based on cellular localisation and functional characteristics. SENP1 and SENP2, form the first subfamily, observed to be mainly nuclear in localisation, and concentrated in the nuclear pore complex (Gong et al., 2000; Bailey & O'Hare 2004). However, due to the presence of nuclear export signals SENP1 and SENP2 can shuttle in and out of the nucleus (Itahana et al 2006). SENP1 and SENP2 have been shown to actively deconjugate SUMO from a wide variety of substrates and have also been observed to have a role in SUMO processing; hydrolysing pre-SUMO1 and pre-SUMO2 into the respective

mature forms, unveiling the C-terminal Gly-Gly motif (Xu and Au, 2005). SENP1 has been shown to be most efficient at processing SUMO-1 to its mature form, whereas SENP2 has been shown to be most efficient at processing SUMO2 (Sharma et al., 2013). *In vitro*, both SENP1 and SENP2 are highly efficient at deconjugating both SUMO1 and SUMO2 (Reverter & Lima 2006; Shen et al., 2006B). SENP3 and SENP5 form the second subfamily of the SENP enzymes and are mainly located in the nucleolus, where they have a role in ribosome biogenesis. SENP3 and SENP5 act to deconjugate SUMO2, showing little activity against SUMO1 (Gong & Yeh 2006). SENP5, like SENP1 and SENP2, can process pre-SUMO2 into its mature form. SENP6 and SENP7 form the third SENP subfamily group, located in the nucleoplasm. These enzymes show weak SUMO processing activity but are active SUMO deconjugating enzymes, thought to mainly depolymerise SUMO chains. In this respect, SENP6 and SENP7 are thought to be SUMO chain editing enzymes (Mukhopadhyay et al 2006; Shen et al., 2009; Hattersley et al., 2011).

The cytoplasmic DeSUMOylating isopeptidases were first identified in 2012 as active SUMO deconjugating enzymes, with very limited SUMO processing activity (Suh et al., 2012). Two DeSUMOylating isopeptidase have been identified; Desi-1 and Desi-2 and are characterised by the presence of permutated papin fold peptidase of the double-stranded RNA viruses and eukaryotes (PPPDE) domain (Suh et al., 2012; Shin et al., 2012). Interestingly, Desi-1 is thought to be active as a homodimer, where the active site is positioned in the groove between two proteins forming a catalytic dyad with two conserved cysteine and two histidine residues (Suh et al., 2012). Unlike the SENP enzymes, Desi-1 and Desi-2 are concentrated in the cytoplasm. Desi-1 and Desi-2 appear to be highly specific enzymes functioning on a very limited substrate selection (Shin et al., 2012).

The most recently identified deSUMOylating enzyme is USPL1, identified in late 2012 (Schulz et al., 2012). USPL1 is distinct from both the SENP and Desi proteins, but interestingly shares around 20% sequence homology with the ubiquitin specific protease USP1 (Schulz et al., 2012).

USPL1 is active in both SUMO processing and SUMO deconjugation but no substrates have yet been identified *in vivo* (Schulz et al., 2012; Hutten et al., 2014).

1.3 Aims and objectives

Initial evidence suggested that the SUMO and ubiquitin conjugation systems were distinct from one and other. However, recent evidence has suggested points of convergence in the two systems. The aim of this thesis is to explore the overlapping ground between the two systems by investigating the notion that polymeric chains composed of both SUMO and ubiquitin can act as unique signalling mechanisms within the cell. These SUMO-ubiquitin hybrid chains may be distinct from the homotypic signals generated by the SUMO and ubiquitin moieties that make up these chains. If so, proteins containing both SUMO and ubiquitin interacting elements should recognise SUMO-ubiquitin hybrid chains with greater affinity than that of SUMO or ubiquitin chains alone.

2 Materials and Methods

2.1 Materials

Information on materials can be found in Appendix I.

2.2 Methods

2.2.1 General Methods

2.2.1.1 General PCR procedure for Phusion™ Hot Start polymerase

List of primers- **2.1.3 in Appendix I**
PCR reaction contents;

DNA template	20-50 ng
5X buffer (HF or GC)	10 µl
10 mM dNTPs	1 µl
10 µM forward primer	1 µl
10 µM reverse primer	1 µl
Phusion polymerase	0.5 µl
dH_2O	to total 50 µl

PCR-program;

Phase	Temp	Duration
Denaturation	98°C	30s
Denaturation*	98°C	10s
Annealing*	lowest primer -2°C	30s
Elongation*	72°C	15s per kb
**cycle 33 times*		
Final Elongation	72°C	10 minutes
Cooling	4°C	∞

PCR product was then mixed with 6X DNA loading buffer and verified on a 0.8% agarose gel. Positive PCR products were then extracted using QIAquick Gel Extraction Kit (Qiagen) following manufacturers protocol.

3.2.1.2 General PCR procedure for site-directed mutagenesis with KOD polymerase

Primer design required mutations with ~15 bases on either side. List of primers- **2.1.3 in Appendix I**

PCR reaction contents;

DNA template	20-50 ng
10X KOD buffer	5 µl
2 mM dNTPs	5 µl
10 µM forward primer	1.25 µl
10 µM reverse primer	1.25 µl
$MgSO_4$	3 µl
KOD polymerase	1 µl
dH_2O	add to total 50 µl

PCR-program;

Phase	Temp	Duration
Denaturation	95°C	2 minutes
Denaturation*	95°C	20s
Annealing*	Primer specific	30s
Elongation*	70°C	1 minutes per kb
**cycle 18 times*		
Final Elongation	70°C	5 minutes
Cooling	4°C	∞

Following PCR reactions 1µl DpnI was added and incubated at 37°C for 1 hour to digest the parental methylated DNA before transformation into DH5α E. coli cells.

2.2.1.3 Restriction digestion and ligation

Plasmid DNA or PCR products were digested with restriction and ligation enzymes from New England Biolabs with the following procedure;

Plasmid DNA (500ng) or 50% of PCR product was incubated at 37°C for 2 hours with 1µl each of the two required restriction enzymes and 2.5µl of 10X NEB digestion buffer. H_2O was added to total reaction volume of 25µl. Restriction digestion was then verified on a 0.8% agarose gel.

Approximately 50µg of digested plasmid vector was added to between 5µl and 10µl of digested PCR product insert. Vector to insert ratio was approximately 1:3. Vector and insert were then incubated with 2.5µl T4 ligation buffer and 1µl T4 ligase at room temperature for 2 hours.

2.2.1.4 Transformation of competent E. coli cells

Freshly thawed competent *E. coli* cells (50µl) were incubated with 50µg of plasmid DNA or 5-10µl of ligation mix on ice for 20 minutes. For preparation of plasmid DNA, DH5α E. coli cells were used. For protein expression BL21 (DE3), Rosetta2 (DE3) or BL21 (DE3) (ELaD) strains were used. The competent bacteria/plasmid DNA mix were then heat shocked for 45 seconds at 42°C. Cells were then placed back on ice for 2 minutes before being transferred to a 15ml falcon tube containing 500µl of LB medium and incubated at 37°C, with shaking, for a minimum of 45 minutes. Cell suspension (250µl) was then spread on agar plates containing the required antibiotic resistance (100µg/ml ampicillin, 50µg/ml kanamycin, 35µg/ml chloramphenicol) and incubated at 37°C overnight (~16 hours).

2.2.1.5 Plasmid DNA preparation

5ml of LB medium with required antibiotic resistance was inoculated with a single colony from an agar plate and incubated overnight (~16 hours) at 37°C with shaking. The cultured DH5α cells were then pelleted by centrifugation at 6800g. Plasmid DNA was then purified from the resulting cell pellet using QIAprep Spin Miniprep Kit (Qiagen) following the manufacturers protocol.

2.2.1.6 DNA concentration measurements

DNA concentration was measured on a NanoDrop 2000 spectrophotometer by UV absorbance at 260 nm. NanoDrop was blanked with nuclease-free water, all plasmids were stored in nuclease-free water.

2.2.1.7 DNA sequencing

Plasmid DNA sequencing was carried out by the DNA Sequencing Service, University of Dundee.

2.2.1.8 Protein concentration measurements

Purified recombinant protein concentrations were measured using NanoDrop 2000 spectrophotometer by UV absorbance at 280 nm together with extinction coefficients determined using the ProtParam tool (www.expasy.org/protparam).

Protein concentrations of cell extracts were estimated using DC assay (Bio-Rad) following manufacturers protocol and measured by UV absorbance at 750 nm.

2.2.1.9 Sodium dodecylsulphate-polyacrylamide gel electrophoresis

Prepared samples for separation by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) were incubated at 100°C for 5 minutes in

SDS loading buffer. Samples were loaded and run in either Novex NuPAGE Bis-Tris 10% gels in MOPS (1X) running buffer or 4-12% gels in MES (1X) running buffer. All gels were run at a constant voltage of 180V, for 45-60 minutes.

2.2.1.10 Coomassie Stain/Destain Procedure

Gels containing samples separated by SDS-PAGE were stained in appropriate amount of Coomassie stain for 5-15 minutes with constant rocking. Coomassie was then poured off and gels were washed quickly with a small amount of Destain 1 to remove residual Coomassie stain, which was then discarded. Destain 1 was then reapplied for 30 minutes before being discarded and replaced with Destain 2 for 15 minutes. Application of Destain 2 was repeated every 15 minutes until all excess Coomassie staining was removed.

2.2.1.11 Western Blotting Procedure

Prepared samples denatured in SDS sample buffer were separated by SDS-PAGE in accordance with **2.2.1.9**. Gels were then loaded into western blot sandwich cassettes with polyvinylidene fluoride (PVDF) transfer membranes (Immobilon-P Transfer membrane, Millipore). Before transfer, membranes were incubated in ethanol for 2 minutes. Transfer was performed overnight (12-16 hours) in Tris/Glycine transfer buffer, at 25 mA. Following transfer, membranes were removed and blocked in 3% BSA for 1 hour, then incubated in 3% BSA containing specified primary antibodies of choice for 2 hours at room temperature, or overnight at 4°C. Membranes were then washed 3 times with PBST (1X) and then blotted with species specific horse radish peroxidase (HRP) secondary antibody diluted in 3% BSA for 2 hours at room temperature. Membranes were again washed 3 times in PBST (1X) before application of enhanced chemiluminescence (ECL) solution. Blots were visualised by exposure of X-ray photographic film. Where necessary, membranes were stripped with western blot stripping buffer for 15-30 minutes, at room temperature before being re-probed with relevant primary antibodies.

2.2.2 Protein Expression and purification protocols

2.2.2.1 Expression of His_6-tagged USP2 catalytic domain (Asn259-Met605)

The gene encoding USP2 CaD (Asn258-Met605) was subcloned from pGEX-6p-USP2 (DSTT) and inserted between HindIII/XhoI sites of pHISTEV30a vector. USP2 CaD was expressed as a His_6-tagged fusion protein with a TEV protease cleavage site situated between the His_6-tag and Met1 of USP2 CaD. The protein was expressed in E. coli R2 (DE3) ELaD k/o cells. 10-25 ml of LB medium supplemented with kanamycin (50 µg/ml), chloroamphenicol (35 µg/ml) and apramycin (40 µg/ml) were inoculated with a single colony from a freshly streaked LB plate supplemented with kanamycin, chloroamphenicol and apramycin. The bacterial culture was then incubated overnight at 37°C with shaking at 220 rpm. The following day, 5 ml of overnight culture was used to inoculate 650 ml of LB medium with kanamycin, chloroamphenicol and apramycin in a 2-litre flask. The cell suspensions were then incubated at 37°C with shaking at 220 rpm until OD_{600} reached 0.6-0.8. At this stage, 1 ml of culture was taken as pre induction sample. The 2-litre flasks were then cooled down in ice–cold water bath for 10-20 minutes, followed by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 µM to induce protein expression. Then cell cultures were then incubated at 20°C with shaking at 220 rpm for 17-19 hours until OD_{600} reached >1.8.

2.2.2.2 Purification of His_6-tagged USP2 CaD

Buffers for Ni-NTA chromatography:

Lysis buffer – 50 mM Tris, 150 mM NaCl, 10 mM imidazole, 2 mM benzamidine, cOmplete protease inhibitor cocktail (EDTA-free, Roche), pH 8.0

Binding buffer – 50 mM Tris, 150 mM NaCl, 10 mM imidazole, pH 8.0

Washing buffer – 50 mM Tris, 150 mM NaCl, 250 mM imidazole, 0.5 mM TCEP, pH 8.0

For USP2 CaD, all stages of the protein purification were carried out where possible, at 4°C. Bacterial cells were harvested by centrifugation at 6200 x g for 20 minutes at 4°C. Cell pellets were then gently resuspeneded in 35 ml of lysis buffer per 650 ml of cell culture by shaking. The cell suspension was then flash-frozen in liquid nitrogen, cells could then be stored at -80°C until further use. After thawing, the bacterial cells were lysed by sonication (Digital Sonifier, Branson). The samples were then cleared of insoluble material by centrifugation at 27000 x g for 45 minutes, at 4°C. The supernatant was then filtered through a 0.2 µm filter and loaded onto a column containing Ni-NTA agarose (DSTT), 5 ml of agarose per 650 ml of culture, pre-equilibrated with binding buffer. The column was then washed with 8 column volumes of binding buffer, followed by 8 column volumes of washing buffer to remove any unbound proteins. His_6-USP2 CaD was then eluted from the column in 3 column volumes of elution buffer. To remove the imidazole from the eluted His_6-USP2 CaD, dialysis was performed overnight at 4°C against elution buffer containing no imidazole, with dialysis buffer exchanged 3 times. During the dialysis, the His-tag was removed from the His_6-USP2 CaD fusion protein through the addition of TEV protease at a ratio of 100:1 His_6-USP2 CaD:TEV protease. After cleavage was completed, His_6/USP2 CaD mix was passed again over the Ni-NTA column. Un-tagged USP2 CaD collected from the column flow through was then concentrated in spin concentrators (Viaspin). USP2 CaD was then check for activity, analysed by SDS-PAGE in accordance with **2.2.1.9**, and stored at -80°C until use.

2.2.2.3 Expression and purification of 4xSUMO2 and Ub-4xSUMO2

4xSUMO2 and Ub-4xSUMO2 constructs in pHISTEV30a vector were a kind gift from Dr. A Plechanovova. 4xSUMO2 and Ub-4xSUMO2 were expressed in Rosetta2 (DE3) bacterial cells and purified in accordance with **2.2.2.2**. Purified 4xSUMO2 and Ub-4xSUMO2 were concentrated, analysed by SDS-PAGE in accordance with **2.2.1.9**, and stored at -80°C until use.

2.2.2.4 Expression and purification of Morc3 and Morc3 mutants

The gene encoding full length Morc3 was subcloned from pSC-HA-Morc3 (DSTT) and inserted between HindIII/XhoI sites of pHISTEV30a vector. Morc3 truncated fragment mutants A/B/C/D were subsequently subcloned from pHISTEV30a-Morc3. Site directed mutagenesis with KOD polymerase was utilised to create the Morc3 4xSIM mutant. One SIM region was mutated in each of four sequential rounds of mutagenesis. All Morc3 variants were expressed in Rosetta2 (DE3) bacterial cells and purified in accordance with **2.2.2.2**. Purified Morc3 and Morc3 mutants were concentrated, analysed by SDS-PAGE in accordance with **2.2.1.9**, and stored at -80°C until use.

2.2.2.5 Expression and purification of USP28 and USP28 mutants

The gene encoding full length USP28 was subcloned from pGEX-6p-USP28 (DSTT) and inserted between BamHI/XhoI sites of pHISTEV30a vector. USP28 N-terminal fragments mutants SIM-UIM/SIM-ΔUIM/ΔSIM-UIM were subsequently subcloned from pHISTEV30a-USP28. All USP28 variants were expressed in Rosetta2 (DE3) bacterial cells and purified in accordance with **2.2.2.2**. Purified USP28 and USP28 variants were concentrated, analysed by SDS-PAGE in accordance with **2.2.1.9**, and stored at -80°C until use.

2.2.2.6 Expression and purification of UBE2v2

His_6-MBP-Ube2v2 was a gift from Dr. A Plechanovova. The protein was expressed in E. coli BL21 (DE3) cells. His_6-/MBP-tags were removed from the His_6-MBP-Ube2v2 fusion protein through the addition of TEV protease at a ratio of 100:1 His_6-MBP-Ube2v2:TEV protease. After cleavage was completed, UBE2v2/MBP/His_6 mix was separated by gel filtration on an Akta system using a HiLoad 16/60 Superdex75 column. Peaked fractions were the analysed by SDS-PAGE in accordance with **2.2.1.9** and fractions containing UBE2v2 were pooled, concentrated and stored at -80°C until use.

2.2.2.7 In vitro transcription/translation

^{35}S labelled proteins were generated using in vitro transcription/translation, from 1 µg plasmid DNA and a wheat germ or rabbit reticulocyte transcription/translation system with ^{35}S as follows: -

Wheat germ lysate/rabbit retics	12.5 µl
Reaction buffer (25X)	1 µl
T7 RNA polymerase	0.5 µl
Amino acid mixture minus methionine	0.5 µl
DNA template	1 µg
^{35}S methionine	1.25 µl
Nuclease free water	to total 25 µl

Reactions were incubated at 30°C for 2 hours before 1 µl of each reaction was separated by SDS-PAGE in accordance with **2.2.1.9**. Gels were dehydrated onto 3mm chromatography paper and exposed to imaging plates overnight. Imaging plates were then scanned through a FujiFilm FLA-5100 and analysed using Aida Image Analyzer software.

2.2.3 Biochemical assays

2.2.3.1 In vitro ubiquitylation assay of Ub-4xSUMO2 and 4xSUMO2

35 nM Uba1 (E1), 35 nM Ubc13/Ube2v2 (E2 pair), 35 nM RNF4 (E3) and 10 µM ubiquitin were incubated together with 2 µM 4xSUMO2 or Ub-4xSUMO2 in 50 mM Tris, 200 mM NaCl, 5 mM $MgCl_2$, 3 mM ATP, 0.1% NP-40 at 37°C. Samples were taken at several different time points, and reactions were stopped via the addition of 2x SDS-PAGE loading buffer. Samples were analysed by SDS-PAGE in accordance with **2.2.1.9**.

2.2.3.2 In vitro deubiquitylation assays for recombinant proteins

0.01, 0.05, 0.1, 0.5, 1 and 5 μM concentrations of deubiquitinases; USP2 CaD or USP28 were incubated with 5 μM substrate protein; 4xSUMO2, Ub-4xSUMO2, K63 pUB or K48 triUb for 4 hours. Reactions were carried out in 50 mM Tris, 150 mM NaCl, 0.05 mM TCEP. Samples were taken at several time points, reactions were stopped by the addition of 2xSDS-PAGE loading buffer at 4 hours, and analysed by SDS-PAGE in accordance with **2.2.1.9**. For control sample USP2 or USP28 were absent from reaction.

2.2.3.3 In vitro deSUMOylation assays

0.01, 0.05, 0.1, 0.5, 1 and 5 μM concentrations of sentrin-specific protease SENP1 (Dr L. Shen) were incubated with 5 μM substrate proteins; 4xSUMO2, Ub-4xSUMO2 and K63 pUB or K48 triUb for 4 hours. Reactions were carried out in 50 mM Tris, 150 mM NaCl, 0.05 mM TCEP. Samples were taken at several time points, reactions were stopped by the addition of 2xSDS-PAGE loading buffer at 4 hours and analysed by SDS-PAGE. For control samples SENP1 was absent from the reaction.

2.2.3.4 TCA precipitation

Sample proteins were precipitated in 20% TCA of equal volume to sample and incubated on ice for 30 minutes. TCA/sample mixes were then centrifuged at 17000 x g for 30 minutes, at 4°C. Supernatants were carefully removed from pellets. 100% ethanol was added to the pellets, mixed by inversion and then centrifuged at 17000 x g for 30 minutes at 4°C. Ethanol was then removed and pellets were dried before resuspension in 2x SDS sample buffer.

2.2.4 Pull-down experiments

2.2.4.1 Linkage of recombinant SUMO, ubiquitin and SUMO ubiquitin hybrid chains to NHS-activated resin

4xSUMO2, K63 pUb, Ub-4xSUMO2, pUb-4xSUMO2 protein were dialysied against 0.2M $NaHCO_3$, 0.5M NaCl, pH 8.3 overnight, at 4°C with two changes of dialysis buffer. Proteins concentrations were then adjusted to equimolar concentrations (>1 mg/ml). N-hydroxysuccinimide (NHS) resin (GE Healthcare) was activated by rinsing with 1 mM HCl before washing with 0.2M $NaHCO_3$, 0.5M NaCl, pH 8.3. For each chain type, 1 ml of NHS activated resin (50% resin slurry in 0.2M $NaHCO_3$, 0.5M NaCl, pH 8.3) was incubated by rotation with 1 ml of protein overnight, at 4°C. A blank resin control was also produced with no protein bound to activated resin. Any free active sites on NHS beads were blocked by sequential incubation with 0.5 M NaCl, 0.5 M Ethanolamine, pH 8.3 followed by 0.5 M NaCl, 0.1 NaOAc, pH 4. Protein bound resins were stored as a 50% slurry in PBS, 0.5 M NaCl, 0.1% NaN_3^-, at 4°C.

2.2.4.2 SUMO, ubiquitin, and SUMO ubiquitin hybrid chain type resin pull-downs for recombinant proteins

The following resin types (50 µl of each); blank, 4xSUMO2, K63 pUb, Ub-4xSUMO2, pUb-4xSUMO2 were first washed with Ultrapure H_2O before equilibration in binding buffer (50 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 0.05% NP-40). 10 µM of recombinant protein was incubated separately with 50 µl of each resin type overnight, at 4°C. After incubation, resins were collected by centrifugation at 500 x g for 1 minute. Supernatants were collected. Resins were then washed once with binding buffer before being transferred to fresh Eppendorf tubes and washed again. Bound proteins were eluted with SDS-loading buffer and analysed by SDS-PAGE and western blot in accordance with **2.2.1.10**. SDS-PAGE gels were imaged with Bio-Rad chemidoc MP imaging system and quantified using ImageJ.

2.2.4.3 SENP1 and USP2 Elution of proteins from SUMO, ubiquitin, and SUMO ubiquitin hybrid chain resin types

After either blank, pUb K63, 4xSUMO2, Ub-4xSUMO2, pUb-4xSUMO2 resins were incubated with substrates and following washing steps, 5 µM of USP2 CaD was incubated with each resin at room temperature for 2 hours. USP2 CaD containing supernatant was removed and mixed with 6X SDS loading buffer. 5 µM of SENP1 was then incubated with each resin at room temperature for 2 hours. SENP1 containing supernatant was then removed and mixed with 6X SDS loading buffer. The five resins types were finally incubated with 2X SDS loading buffer for 5 minutes at room temperature, then at 100°C for 10 minutes. Samples for each of the five resin types and for the separate elutions were then separated by SDS-PAGE before analysis by western blot.

2.2.5 Cell Culture

2.2.5.1 Cell lines

All cells, unless stated, were obtained from Hay lab stocks, and maintained at 37°C and 5% CO_2.

Wild type Hela, U20S, HCT116 and 293 N3S cells were cultured in monolayers maintained in Dulbecco's Modified Eagle Medium plus Glutamax (DMEM) (Invitrogen) supplemented with 10% fetal calf serum (FCS) and 100U/ml penicillin and streptomycin (Invitrogen).

293 N3S cells cultured in suspension were maintained in Minimum Essential Eagles Medium Spinner Modification (SMEM) (Sigma) supplemented with L-glutamine, 10% FCS and 100U/ml penicillin and streptomycin. Cells were maintained at 37°C under atmospheric CO_2.

U2OS ΔRNF4 and HCT 116 ΔRNF4 clones, a kind gift from Dr Jean-François Maure (Hay Lab), were cultured in monolayers maintained in Dulbecco's

Modified Eagle Medium plus Glutamax supplemented with 10% FCS and 100U/ml penicillin and streptomycin.

U2OS cells expressing GFP-RAP80 (kind gift from Prof. SP Jackson) were cultured in monolayers maintained in Dulbecco's Modified Eagle Medium plus Glutamax supplemented with 10% FCS and 100 U/ml penicillin and streptomycin. Cells were maintained under antibiotic selection with 100 mM puromycin.

2.2.5.2 Cell extract preparation – for the preservation of SUMO and ubiquitin modification of proteins

Buffers for cell extract preparation:
Hypotonic buffer A -- 10 mM HEPES pH7.9, 1.5 mM $MgCl_2$, 10 mM KCl_2, 0.08% NP-40, Roche cOmplete, EDTA-free protease inhibitor tablets, 20 mM iodoacetamide.
RIPA (5X) -- 250 mM Tris pH 7.5, 750 mM NaCl, 5% NP-40, 2.5% DoC, Roche cOmplete, EDTA-free protease inhibitor tablets, 100 mM iodoacetamide.

RIPA (1X) -- 50 mM Tris pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% DoC, Roche cOmplete, EDTA-free protease inhibitor tablets, 20 mM iodoacetamide.

Cells were washed 3 times in ice cold PBS containing 200 mM iodoacetamide and pelleted. Cells were then resuspended in ice cold hypotonic buffer A to a final volume of approximately 10x the cell pellet and left on ice for 15 minutes. Cells were then lysed by repeatedly being forced through a 25-guage needle. Resulting cell lysates were then spun at 500 x g for 5 minutes at 4°C to pellet nuclei. Supernatant was removed. RIPA buffer (5X) was then added to the supernatant in a ratio of 1:4, resulting in the cytoplasmic fraction. The nuclei-containing pellet was then washed once in hypotonic buffer A, resuspended in RIPA buffer (1X) and sonicated low setting (Diagenode, Biorupter) on ice until the nuclear membrane was disrupted. Benzonase 1 Unit (Merek) was added to disrupted nuclear extracts for 30 mins on ice to digest any remaining DNA.

Both cytoplasmic and nuclear samples were then cleared by centrifugation at 17000 x g for 15 minutes. Protein concentrations of nuclear and cytoplasmic extracts were estimated using Bio-Rad DC Assay.

2.2.5.3 Whole Cell Extract Preparation

Cells were washed 3 times with ice cold PBS, scraped in an appropriate volume of ice cold PBS and then pelleted by centrifugation at 500 x g. Cells were then lysed in 2X SDS lysis buffer and sonicated at high setting for 180 seconds (Diagenode, Biorupter). Samples were then incubated at 100°C for 10 minutes before being cleared by centrifugation at 17,000g for 10 minutes. Protein concentration was estimated using Bio-Rad DC Assay. B-mercaptoethanol was then added to a final concentration of 200 mM before samples were separated by SDS-PAGE for western blotting analysis in accordance with in accordance with **2.2.1.10**.

2.2.5.4 SiRNA Transfections

Cells were seeded to be 60% confluent at transfection, one day prior to transfection. For each well of a 6 well plate, 1.25 µl of RNAiMAX (Invitrogen) was incubated with 100 µl of opti-MEM serum free medium (Invitrogen) for 5 minutes. Separately, 0.25 µl of siRNA 20 µM was incubated with 100 µl of opti-MEM serum free medium for 5 minutes. RNAiMAX and siRNA containing opti-MEM solutions were then combined, gently mixed and incubated for 15 minutes at room temperature. RNAiMAX/siRNA containing opti-MEM solution was then added to cells in 0.8 mL of DMEM containing 10% FBS for a final siRNA concentration of 10 nM. Cells were then incubated at 37°C for 5-6 hours before DMEM containing RNAiMAX/siRNA opti-MEM solution was removed and cells were incubated with DMEM supplemented with 10% FBS 100U/mL penicillin and streptomycin and incubated for 48 hours prior to analysis.

2.2.5.5 Cell cycle analysis by flow cytometry

HCT116 cells were harvested by standard techniques and resuspended in 1ml of PBS supplemented with 1% FCS. Cells were then transferred to FACS tubes and pelleted. PBS was removed and cells were fixed in 1 ml of ice cold 70% EtOH for 30 minutes at room temperature with gentle vortexing to prevent clumping. Cells were adjusted to 5 x 10^5 cells/ml and washed twice in PBS supplemented with 1% FCS. Cells were then incubated in 300µl of FACS staining buffer, protected from light, for 20 minutes, at room temperature. Cells were then analysed on a FACS Calibur Analyser by Dr R. Clark.

2.2.6 Immunoprecipitation (IP)

2.2.6.1 USP28 & Morc3 IPs

Hela cells were seeded on ten 10 cm dishes and grown until 100% confluent. Nuclear extracts were generated as described in **2.2.5.3**. 5 mg of whole cell extracts was precleared by incubation with 50 µl of Protein G Dynabeads (Invitrogen) for 1 hour with continuous rolling. For each IP conditon, 2 µg of appropriate antibodies were incubated with 50 µl of Protein G Dynabeads. 5 mg of whole cell extracts were then incubated with 50 µl of antibody-bound Dynabeads overnight at 4°C with continuous rolling. Dynabeads were then removed from cell extracts and washed 3 times with RIPA buffer (1X). Bound proteins were eluted with 2xLDS sample buffer and analysed via western blotting in accordance with **2.2.1.10**.

2.2.6.2 RAP80 IP

For each treatment, 293 N3S cells were seeded on five 10 cm dishes and grown until 100% confluent. Nuclear extracts were generated as described in **2.2.5.3**. 2.5 mg of nuclear extract was precleared by incubation with 50 µl of Protein G Dynabeads (Invitrogen) for 1 hour with continuous rolling. For each IP conditon, 2 µg RAP80 antibodies were crosslinked to 50 µl of Protein G

Dynabeads with 5 μM BS^3 crosslinking agent. For SENP1 treatment, nuclear extracts were incubated with 20 μM SENP1 for 4 hours, at 4°C prior to incubation with RAP80 antibody-bound Dynabeads. 2 mg of protein extract were then incubated with 50 μl of antibody-bound Dynabeads overnight at 4°C with continuous rolling. Dynabeads were then removed from cell extracts and washed 3 times with RIPA buffer (1X). Bound proteins were then eluted with 2xLDS sample buffer and analysed via western blotting in accordance with **2.2.1.10**.

2.2.7 Microscopy

2.2.7.1 Adherent cell culture on coverslips

Cells were cultured on 10 mm glass coverslips in DMEM supplemented with 10% FCS plus pen/strep until 60% confluent in 30 mm dishes. Cells were then irradiated if required and allowed to recover for specified time. Cells were washed 3 times with ice cold PBS and then fixed to the coverslips with 2% paraformaldehyde (PFA) for 10 minutes. Cells were then washed 3 times with PBS and permeabilised by incubation with 0.5% NP-40 in PBS for 10 minutes followed by one wash with PBST. Cells were incubated with blocking agent 3% BSA/PBST, at room temperature with shaking for 30-90 minutes, or overnight at 4°C. Cells were then washed with PBST 3 times. Coverslips were removed from 30 mm dishes and place on parafilm layered in a wet chamber.

2.2.7.2 Immunofluorescence

Adherent cells were processed as described in **2.2.7.1.** Primary antibodies diluted in blocking buffer were then incubated on coverslips in a wet chamber for 60 minutes at room temperature, followed by 3 washes with PBST. Away from light sources, secondary Alexa fluor antibodies diluted in blocking buffer were then incubated on coverslips for 60 minutes at room temperature, followed by 3 washes with PBST. Cells were stained with 1μg/ml DAPI for 5 minutes. Cells were then washed 3 times with PBST followed by one wash

with PBS. Any residual liquid was aspirated from coverslips. Coverslips were mounted to glass microscope slides using Vectashield mounting media (Vector labs), and sealed down with nail varnish. Images were then captured using a DeltaVision Elite microscope (Applied Precision).

2.2.7.3 In situ Proximity Ligation Assays (PLA)

Adherent cells were first processed as described **2.2.7.1.** All PLA steps for 10 mm coverslips were carried out in a minimum of 40 µl to ensure total coverage of coverslip. Firstly, Primary antibodies were diluted in blocking buffer and incubated on coverslips in a wet chamber for 2 hours at room temperature, or overnight at 4°C. Primary antibody solutions were then removed by aspiration and followed by two 5 minute PBST washes with gentle shaking. For each coverslip, 8 µl of PLA probe MINUS stock, 8 µl of PLA probe PLUS stock and 24 µl of Antibody Diluent were mixed together before incubation on coverslips for 1 hour at 37°C in a wet chamber. PLA probe solutions were then removed by aspiration and coverslips were washed three times in PBST for 5 minutes with gentle shaking. Ligation of the PLA probes was then carried out. 8 µl of 5x Ligation stock was added to 31 µl of nuclease-free H_2O resulting in ligase solution. Immediately prior to incubation with coverslips 1 µl of PLA ligase was added to 39 µl of ligase solution, pipetted up and down 5 times and added to coverslips. Coverslips were incubated in a pre-heated wet chamber for 30 minutes at 37°C. Ligation-ligase solution was then removed by aspiration and two 1 minutes washes with PBST were carried out. Amplification of the PLA signal was then carried out requiring minimal exposure to light for all further experimental steps. 8 µl of 5x Amplification stock and 31.5 µl of nuclease-free water were mixed with 0.5 µl of Polymerase and then added to the coverslips. Coverslips were then incubated in a preheated wet chamber for 100 minutes at 37°C. Polymerase solution was then aspirated off coverslips. Coverslips were washed twice in 1x PBST for 10 minutes with shaking followed by one 1 minute wash with PBS. Coverslips were allowed to dry for 15 minutes at room temperature protected from light. Coverslips were then mounted to glass microscope slides using Vectashield mounting media and sealed down with nail varnish.

Images were then captured using either Deltavision or spinning disk confocal microscopes.

2.2.7.4 Imaging

Deltavision – Immunofluorescence samples were imaged by widefield microscopy using a DeltaVision Elite microscope (Applied Precision). Typically, images of multiple z-sections were captured using a coolsnap HQ CCD camera and 60x oil immersion lenses. (Centre for Advanced Scientific Technologies, Dundee)

Spinning disk confocal – a subset of PLA assay samples were imaged using a MAG Biosystems EnVision spinning disk microscope equipped with 440 nm, 491 nm and 594 nm laser lines. (Centre for Advanced Scientific Technologies, Dundee)

2.2.7.5 Image analysis

Images were deconvoluted using the spinlock deconvolution cluster (spinlock.openmicroscopy.org.uk). Deconvolved images were then analysed using OMERO software (OME Open Microscope Environment). Foci counting and analysis was performed using mtools plug-ins developed by Mr Michael Porter. Foci were determine as pixel clusters above a given fluorescence threshold in maximal intensity projections of multiple Z- sections, and this threshold was maintained for the analysis of all images in a given experiment.

2.2.8 Mass spectrometry

2.2.8.1 In-gel trypsin digestion

For all mass spectrometry analysis, microwave assisted in-gel trypsin digestion of proteins was utilised. Microwave steps were performed while tubes were held in a proprietary container that suspended samples in 1.5 L water during irradiation. Water was changed once temperature was greater

than 55°C. All steps involving the addition or removal of solutions were carried out in a fume hood with fresh filter tips used at every stage.

All protein samples for proteomics were run on NuPAGE Novex 10% Bis-Tris in (1X) MOPS buffer and stained/destained according to Coomassie stain protocol. After destaining to the desired level, gels were washed with 15 ml of pure H_2O for at least 30 minutes. Gels were then cut as appropriate into slices. Slices were then diced into ~1 mm cubes and stored in Eppendorf LoBind tubes at 4°C until processed.

Gel slices were then further destained by microwaving on full power (700 Watts) in 500-900 µl 50 mM ammonium bicarbonate (ABC), 50% acetonitrile (ACN) until all blue stain was removed from slices. Slices were then spun briefly and liquid was removed from gel pieces. Gel pieces were then dehydrated by adding 200-700 µl 100% ACN with vortexing, liquid was then removed and gel pieces were allowed to dry for 10-20 minutes at room temperature in a fume hood.

Disulfide adducts were then reduced through incubation with 500 µl of 10 mM DTT in 100 mM ABC with microwaving at (420 watts) for 10 minutes. DTT ABC solution was then removed from gel pieces. Proteins were then alkylated through the addition of 300-700 µl 50 mM iodoacetamide in 100 mM ABC with shaking, protected from light, at room temperature. Gel pieces were then washed sequentially at room temperature, with shaking for 15 minutes with; 400-900 µl 100 mM ABC; then 400-900 µl of 200 mM ABC, 50% ACN and finally dehydrated with 300-700 µl of 100% ACN with vortexing. All residual liquid was then removed and gel pieces were dried for 10-20 minutes. Per lane, assuming <25 µg of total protein in that lane, 1.25 µg of trypsin was diluted in 400-600 µl 20 mM ABC, 9% ACN. Trypsin solution was then added evenly to gel pieces until fully rehydrated. Further 20 mM ABC, 9% ACN solution was added until gel pieces in each tube were completely covered by liquid. Trypsin solution and gel pieces were then incubated for 12-16 hours at 37°C.

Peptides were then extracted by the addition of 100% ACN equal to the amount of trypsin solution in each tube and incubation at room temperature for 30 minutes, with vortexing. Liquid was then transferred to a new Eppendorf tube. 5% formic acid, 50% ACN was then added to the gel

pieces, incubated for 30 minutes with vortexing then transferred to the Eppendorf tube containing trypsin ACN solution. Finally the gel pieces were dehybdrated with 50-100 µl of 100% ACN. Incubated for 10 minutes with vortexing then transferred to the Eppendorf containing trypsin ACN solution.
The peptide-containing trypsin solution was then evaporated by gryovac and the peptide pellet resuspended in 35 µl of 0.5% acetic acid/0.1% TFA ready to be analysed.

2.2.8.2 Non-labelled hybrid chain resins pull-down experiment

293 N3S cells were initially cultures as monolayers maintained in Dublecco's Modified Eagle Medium plus Glutamax (DMEM) (Invitrogen) supplemented with 10% FCS and 100 U/ml penicillin and streptomycin. Cells were grown in three 150 cm^2 flasks to 90% confluency and then transferred into 500 ml of Minimum Essential Eagles Medium Spinner Modification (SMEM) (Sigma) supplemented with L-glutamine, 10% FCS and 100 U/ml penicillin and streptomycin. Cells were maintained at 37°C in a sealed 1 litre Duran under non-adjusted CO_2 and expanded to a total of 45 litres of cells at a density of ~1 $x10^6$ cells/ml. Nuclear cell extracts were then isolated as described in **2.2.5.3**. Blank, K63 pUb, 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 bound NHS resins were then each separately incubated with 10 ml of nuclear lysate (5.5 mg/ml) overnight at 4°C with rotation. Resins were washed 3 times with RIPA (1X) buffer before SENP1 and USP2 CaD elution was performed as described in **2.2.4.3.** Samples for each elution were separated by SDS-PAGE. Elutions were each run on a single lane. Each lane was cut into 2 slices (30 in total). In-gel trypsin digestion was performed in accordance with **2.2.8.1.** 8 µl of each slice was run on a 90 minute HPLC gradient, Easyspray 45 cm column. Samples were analysed by liquid chromatography-tandem MS on a Q Exactive mass spectrometer (Thermo Scientific). RAW MS data files were processed using MaxQuant software (version 1.3.0.5) and searched against the UniProtKB human proteome.

2.2.8.3 RAP80 IP of SENP1 treated extract samples

For each treatment, 293 N3S cells were seeded on five 10cm dishes and

grown until 100% confluent. For ionising radiation treatment (IR), Cells were treated with 4 Gys of Y-radiation and left to recover at 37°C for 1 hour. Nuclear extracts were generated as described in **2.2.5.3**. 2.5 mg (1ml) of nuclear extracts were precleared by incubation with 50µl of Protein G Dynabeads (Invitrogen) for 1 hour with continuous rolling. For each IP conditon, 2 µg RAP80 antibodies were crosslinked to 50 µl of Protein G Dynabeads with 5 µM BS^3 crosslinking agent. For SENP1 treatment, nuclear extracts were incubated with 20 µM SENP1 for 4 hours, at 4°C prior to incubation with RAP80 antibody-bound Dynabeads. 2 mg of protein extract were then incubated with 50 µl of antibody-bound Dynabeads overnight at 4°C with continuous rolling. Dynabeads were then removed from cell extracts and washed 3 times with RIPA buffer (1X). Bound proteins were then eluted with 2xLDS sample buffer and analysed via western blot. Samples were run on SDS-PAGE and analysed via mass spectrometry. In-gel trypsin digestion was performed in accordance with **2.2.8.1**. Approximately 1 µg of each slice was run on a 150 minute HPLC gradient, Easyspray 45 cm column. Samples were analysed by liquid chromatography-tandem MS on a Q Exactive mass spectrometer (Thermo Scientific). RAW MS data files were processed using MaxQuant software (version 1.3.0.5) and searched against the UniProtKB human proteome.

2.2.8.4 RAP80 IP of SILAC labelled cell extracts with USP2 CaD/SENP1 elution

293 N3S cells were initially cultured as monolayers maintained in Dulbecco's Modified Eagle Medium plus Glutamax (DMEM) (Invitrogen) supplemented with 10% FCS and 100U/ml penicillin and streptomycin prior to amino acid labelling. Cells were then spilt into three 75 cm^2 dishes (Light, Medium, Heavy) and cultured in SILAC DMEM supplemented with 10% dialysed FCS and $_L$-lysine and $_L$-arginine replaced with stable isotope SILAC forms (Cambridge Isotope Laboratories). "Light" cells were grown in SILAC DMEM containing isotopically "normal" amino acids, K0R0. "Medium" cells were grown SILAC DMEM containing 4,4,5,5-D_4 lysine and $^{13}C_6$ arginine, K4R6. "Heavy" cells were grown in SILAC DMEM containing $^{13}C_6$ $^{15}N_2$ lysine and

$^{13}C_6{}^{15}N_4$ arginine, K8R10. Light, Medium and Heavy SILAC labelled cells were cultured until 90% confluent in respective SILAC DMEM, then split 1/10. This process was repeated five times to allow all proteins within the cultures to incorporate labelled amino acids. Twenty 10 cm dishes were seeded for each of the three SILAC labelled cell conditions. Cells were grown until 100% confluent irradiated with 4 Gys of γ-irradiation and before nuclear extracts were isolated as described in **2.2.5.3**. RAP80 IPs were performed overnight at 4°C, for each of the three SILAC conditions. Heavy, Medium and Light RAP80 IP bound resins were then washed three times with RIPA (1X) containing no iodoacetamide. Heavy RAP80 IPs were then incubated with 10 μM SENP1 for 4 hours at room temperature. Medium RAP80 IPs were incubated with 10 μM USP2 CaD for 4 hours at room temperature. Light RAP80 IPs were incubated in RIPA (1X) containing no iodoacetamide for 4 hours at room temperature. SENP1, USP2 and RIPA control supernatants were then removed from RAP80 IP resins and TCA precipitations were performed on each supernatant as described in **2.2.3.4**. 80 μl of 2XLDS sample buffer was then incubated sequentially with each of the three RAP80 IP resins before resuspending each of the TCA precipitated SENP1, USP2 and RIPA elutions. 35 μl of the resulting samples were separated by SDS-PAGE. The lane containing separated material was then cut into 24 slices. The molecular weight range for each slice was noted before in-gel trypsin digestion was performed in accordance with **2.2.8.1**. 8 μl of each slice was run on a 150 minute HPLC gradient, Easyspray 45 cm column. Samples were analysed by liquid chromatography-tandem MS on a Q Exactive mass spectrometer (Thermo Scientific). RAW MS data files were processed using MaxQuant software (version 1.3.0.5) and searched against the UniProtKB human proteome.

2.2.8.5 General description of mass spectrometry

Mass spectrometry analysis was performed by benchtop orbitrap mass analyzer (Q Exactive, Thermo Fisher Scientific) equipped with an integrated nano-electrospray (Easy-Spray Thermo Fisher Scientific) coupled to a nano-

uHPLC system (Easy N-LC-1000 Thermo Fisher Scientific). Peptides were fractionated on a 50 cm x 75 µm ID, PepMap RSLC C18, 2 µm reverse phase column run at 45°C, fractionating over 1-4 hour gradients (experiment specific). Data were acquired in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Full scan spectra (m/z 300-1800) were acquired in the orbitrap with resolution R = 70,000 at m/z 200 (after accumulation to target value of 1,000,000 or 60 ms).

Raw MS data files were processed with the quantitative mass spectrometry software MaxQuant (version 1.3.0.5) (Cox et al., 2008). Enzymes specificity was set to trypsin-P. Gly-gly adducts to lysine were searched as variable modification. The data were searched against the Human protein database. Maximum allowed mass deviation was set to 20 parts per million for peptide masses and 0.5 da for MS/MS peaks. The minimum peptide length was set to 7 amino acids and the maximum of two missed cleavages. 1% false discovery rate (FDR) was required at both protein and peptide level. In addition to the FDR threshold, proteins were considered identified, if they had at least 1 razor peptide. Each raw data file was considered as a single ‘experiment’, so protein intensity values based on extracted ion chromatographs (XICs) were reported for each slice sample. These intensities were used as an approximation of relative protein abundance for comparing the same protein among samples. Protein intensities were arimetically converted to ratio values and ratios were normailized to the median ration of the entire group to allow for sample loading errors.

3 Establishment and validation of approaches to identify proteins interacting with SUMO-ubiquitin hybrid chains

3.1 Introduction

Initially the SUMO and ubiquitin modification pathways were thought to be distinct, with the action of SUMO and ubiquitin E3 ligases working antagonistically. However, more recently the identification of ubiquitin E3 ligases that are targeted to SUMO modified proteins has begun to lift the veil on the interplay between the SUMO and ubiquitin modification pathways, revealing a potential new form of signalling motif the SUMO-ubiquitin hybrid chain.

3.1.1 SUMO targeted ubiquitin E3 ligases

SUMO targeted ubiquitin E3 ligases (STUbLs) are a recently identified family of enzymes that are targeted to SUMO modifications via internally located SIM domains that are required for substrate modification by the ubiquitin E3 ligase. STUbLs were first identified in the yeast *S. cerevisiae* with the identification of Uls1 and the heterodimeric complex Slx5-Slx8 (Uzunova et al., 2007). Both Uls1 and Slx5-Slx8 have been shown to contain RING domains, and multiple SIM domains (Uzunova et al., 2007). RNF4 was the first human STUbL identified (Lallemand-Breitenbach et al., 2008; Tatham et al., 2008). RNF4 is a small ~21 kDa protein that belongs to the RING type ubiquitin E3 ligases with a RING domain located at it C-terminus. RNF4 has been observed to catalyse the transfer of ubiquitin from ubiquitin loaded E2 enzymes to substrates only after dimerization of its C-terminal RING domain (Liew et al., 2010; Plechanovova et al., 2011; Plechanovova et al., 2012). In addition to its RING domain, RNF4 contains four N-terminally located SIM domains (Tatham et al., 2008). These SIM domains are a mixture of classical and high fidelity SIM types; SIM 2 and 3 of RNF4 conform to the high fidelity type SIM with SIMs 1 and 4 conforming to the more general classical consensus SIM, this allows RNF4 to bind poly-SUMO2 chain of four or more SUMO moieties with high affinity (**Figure 3.1-1**) (Tatham et al., 2008). This

affinity for poly-SUMO has been observed to be critical for the function of RNF4. *In vivo,* RNF4 has been observed to bind poly-SUMOylated PML and subsequently catalyse the addition of K48 linked poly-ubiquitin chains to the previously poly-SUMOylated PML protein; this poly-SUMO poly-ubiquitin modified PML protein is then degraded by the proteasome (Tatham et al., 2008). This action of RNF4 is dependent on the SIM domains located within RNF4 as mutants of RNF4 containing disrupted SIM domains fail catalyse the ubiquitylation of PML (Tatham et al., 2008). This apparent SUMO dependent activity of RNF4 was further demonstrated in a recent paper by Rojas-Fernandez and colleagues where by it was observed that, at physiologically relevant concentrations, RNF4 is found as a monomer with no ubiquitin E3 ligase activity. The presence of poly-SUMO chains however resulted in the dimerization of RNF4, with this SUMO interacting homodimer of RNF4 functioning as a fully active ubiquitin E3 ligase (Rojas-Fernandez et al., 2014). Thereby further cementing that the ubiquitin E3 ligase activity of RNF4 is dependent on its ability to dimerize in the presence of SUMO.

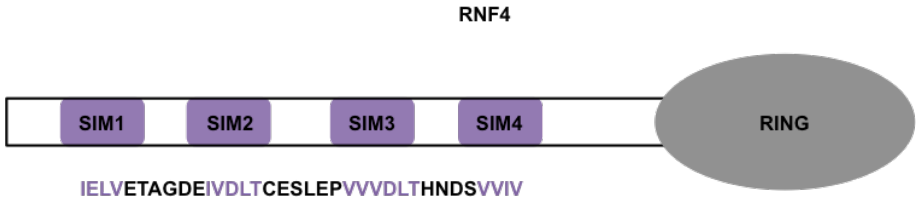


Figure 3.1-1 Domain schematic of RNF4.

RNF4 contains two high fidelity type SIM domains; SIM2 & SIM3 and two general consensus type SIM domains; SIM1 & SIM4. The SIM domains sequences are highlighted in purple. RNF4 also contains a C-terminal RING domain.

More recent work has suggested roles outwith tagging SUMOylated proteins for degradation by the proteasome. RNF4 has been observed to be recruited to SUMO modified mediator of damage checkpoint 1 (MDC1) at sites of DNA damage and play a critical role in the subsequent response to DNA double strand breaks (DSBs) (Galanty et al., 2012; Yin et al., 2012). Depletion of RNF4 in both human and chicken cells shows decrease of K63 linked poly-ubiquitin, but not K48 linked poly-ubiquitin deposited at the sites of DNA damage, an increase in sensitivity to ionising radiation and a reduction in

the efficiency of DNA repair by homologous recombination and non-homologous end joining. This is thought to be due to a failure to load RPA and Rad51 onto DNA at the damage site and thus an inability to efficiently carry out end resection at the DSBs (Yin et al., 2012).

As a result of the action of STUbLs like RNF4, and the more recently described RNF111/Arkadia (Poulsen et al., 2012), hybrid chains of SUMO and ubiquitin are thought to be present at the sites of DNA damage as well as proteins destined for destruction by the proteasome. Thus, these proteins act as points of convergence in the once thought independent SUMOylation and ubiquitylation systems. However, it is still unclear as to the form this convergence takes and whether the signals generated by hybrid chains of SUMO and ubiquitin are distinct to that of homotypic chains of SUMO and ubiquitin modifying the same substrate. Recent evidence has suggested a role for RNF4-dependent SUMO-ubiquitin hybrid chains in the recruitment of the BRCA-1A complex, via the adaptor protein RAP80 to the sites of DNA damage (Guzzo et al., 2012). It was known previously that RAP80 targeted BRCA-1 to the sites of DNA double strand breaks through recognition of K63 linked poly-ubiquitin chains by two UIM domains located in the N-terminal region of RAP80 (Yan et al., 2002). Another key observation is that SUMOylation is required for the recruitment of BRCA-1 to damage sites (Morris et al., 2009). A SIM domain has been identified in close proximity to the two previously characterised UIMs of RAP80, forming a tandem SIM UIMUIM SUMO-ubiquitin binding element that allows RAP80 to bind to SUMO-ubiquitin hybrid chains with higher affinity than either to SUMO or ubiquitin alone (Guzzo et al., 2012) (**Figure 3.1-2**). The discovery of this high affinity binding element in the N-terminal region of RAP80 is intriguing-suggesting that the SUMO-ubiquitin hybrid chain is a novel type of signalling motif distinct to that of separate modifications by homotypic chains of both SUMO and ubiquitin on the same substrate.

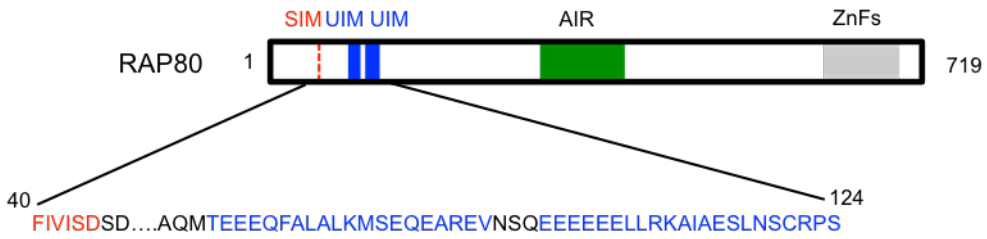


Figure 3.1-2 RAP80 contains a tandem SIM UIM UIM domain.

Schematic representation of BRCA 1A complex component RAP80. RAP80 contains a tandem SIM UIMUIM domain located in its N-termini. Red dashed line represents SIM domain. SIM sequence is highlighted below in red. Blue boxes represent RAP80s two UIM domains. UIM sequences are highlighted in below in blue.

3.1.2 SUMO-ubiquitin hybrid chain assembly

The SUMO-ubiquitin hybrid chain assembly pathway suggests a high level of specificity for this signal. This specificity is due to the multiple SUMO and ubiquitin E1, E2 and E3 enzymes that would be required to conduct the initiation of hybrid chain synthesis and create the signal (**Figure 3.1-3**). The affinity of protein receptors for hybrid chains would also be high due to the requirement of both SUMO and ubiquitin interaction elements, as seen in the tandem SIM UIMUIM of RAP80 (Guzzo et al., 2012).

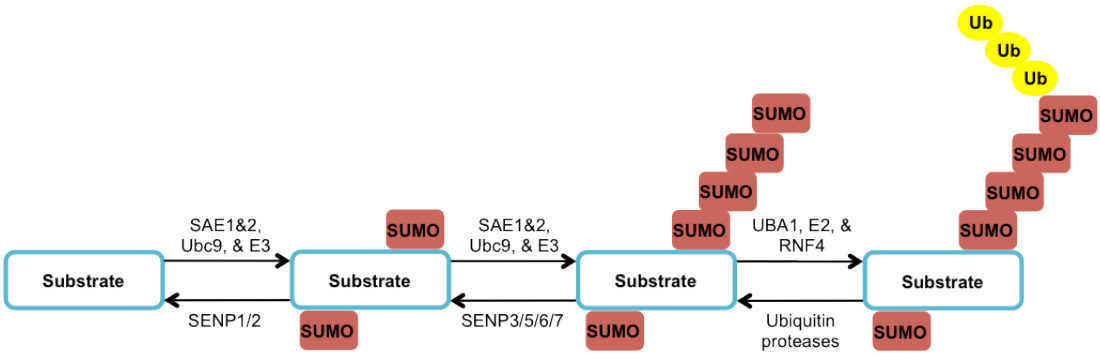


Figure 3.1-3 SUMO-ubiquitin hybrid chain formation pathway.

SUMO-ubiquitin hybrid chain formation requires the orchestration of both the SUMO and ubiquitin conjugating pathways.

The number of potential combinations of covalent interactions between SUMO and ubiquitin is vast. SUMO polymers primarily form on K11 within

SUMO2, however both K5 and K7 of SUMO2 are known to be capable of accepting SUMO modification (Tammsalu et al., 2014). Ubiquitin could potentially modify SUMO2 on any one of eight internal lysine residues, with ubiquitin also known to modify the amino group of the N-terminal Met1 of SUMO2 (Tatham et al., 2013). Ubiquitin chains can then form on any one of seven lysine residues found within ubiquitin or linear chains of ubiquitin could form via Met1. Conversely SUMO is also known to modify ubiquitin thereby creating the possibility of mixed chains of SUMO and ubiquitin **(Figure 3.1-4)** (Tammsalu et al., 2014). Depending on the number of SUMO and ubiquitin moieties involved this could create a huge number of potential combinations of SUMO-ubiquitin hybrid chain, and thus the potential for a great number of very specific signals. However, to date, with the potential exception of RAP80, it is unclear if specific receptors have evolved to recognise SUMO-ubiquitin hybrid chains.

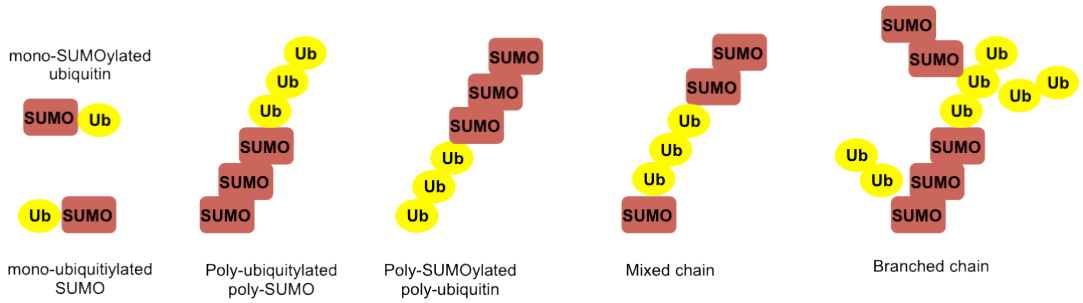


Figure 3.1-4 Possible hybrid chain types.

SUMO is capable of modifying ubiquitin, and ubiquitin is capable of modifying SUMO. In theory this could allow for multiple forms of SUMO-ubiquitin hybrid chain to be formed, including mixed and branched hybrid chain types.

The possibility of SUMO-ubiquitin hybrid chains acting, as a unique signalling device in cells is of interest because of the specificity required to interact with these chains and the affinity that this then suggests. It is interesting to note that STUbLs are not restricted to the mammalian system, and in fact were first identified in yeast (Uzunova et al., 2007). Thus suggesting some evolutionary pressure to conserve mechanisms capable of constructing such signals. RNF4 is known to have many cellular roles and

has the potential to deposit hybrid chains on target substrates in these processes, with the possible result being that SUMO-ubiquitin hybrid chains could play an important role in many cellular processes. Therefore, investigation into the role of SUMO-ubiquitin hybrid chains is worthy of further investigation, and may reveal some important biological functions at the crossover between the SUMO and ubiquitin modification pathways.

3.2 Results

Initial experiments were designed to elucidate a role for RNF4 generated hybrid chains of SUMO and ubiquitin by identifying novel proteins that interact non-covalently, yet specifically to polymeric chains that contain ubiquitin covalently linked to SUMO known as SUMO-ubiquitin hybrid chains. Here, a mass spectrometry based approach was used to identify a wide variety of proteins that appear to show affinity for 'baits' consisting of recombinant protein chains of poly-SUMO, poly-ubiquitin, or hybrid chains of poly-SUMO mono-ubiquitin or poly-SUMO poly-ubiquitin.

3.2.1 Generating SUMO-ubiquitin hybrid chains.

The ability to generate pure hybrid chains consisting of both SUMO and ubiquitin is critical to allow the preparation of affinity tools as bait for proteomic studies. In pursuit of this goal, recombinant chains of poly-SUMO and hybrid chains of SUMO and ubiquitin were generated (**Figure 3.2-1**). Recombinant linear HisSUMO2-3xSUMO2ΔN11 (referred to as 4xSUMO2) (**Figure 3.2-1 A, upper**) and HisUb-SUMO2-3xSUMO2ΔN11 (referred to as Ub-4xSUMO2) (**Figure 3.2-1 A, lower**) were expressed in BL21 (DE3) bacterial cells and purified firstly via Ni-NTA agarose, before a secondary size exclusion chromatograph purification through Superdex75 columns (**Figure 3.2-1 B**). The resulting purified recombinant proteins represented a simplistic, yet biologically relevant (Tatham et al., 2013), linear head-to-tail poly-SUMO chain (4xSUMO2), and a hybrid chain consisting of a linear head-to-tail poly-SUMO chain N-terminally modified with one ubiquitin (Ub-4xSUMO). To generate a poly-SUMO poly-ubiquitin hybrid chain, 4xSUMO2 and Ub-4xSUMO2 were used as substrates in an *in vitro* ubiquitylation assay (**Figure 3.2-1 C**). The K63-specific ubiquitin conjugating E2 pair Ubc13/UBE2v2 (Mckenna et al., 2003) were used in conjunction with the SUMO-targeted ubiquitin E3 ligase RNF4 to generate K63 linked poly-ubiquitin chains. 4xSUMO2 was unable to be ubiquitylated in this assay, with only unanchored ubiquitin chains formed. However, the ubiquitin containing Ub-4xSUMO2 substrate was efficiently ubiquitylated after 60 minutes, as poly-ubiquitin

chains were formed on the acceptor lysines found on the N-terminally located ubiquitin moiety (**Figure 3.2-1 C, left panel**).

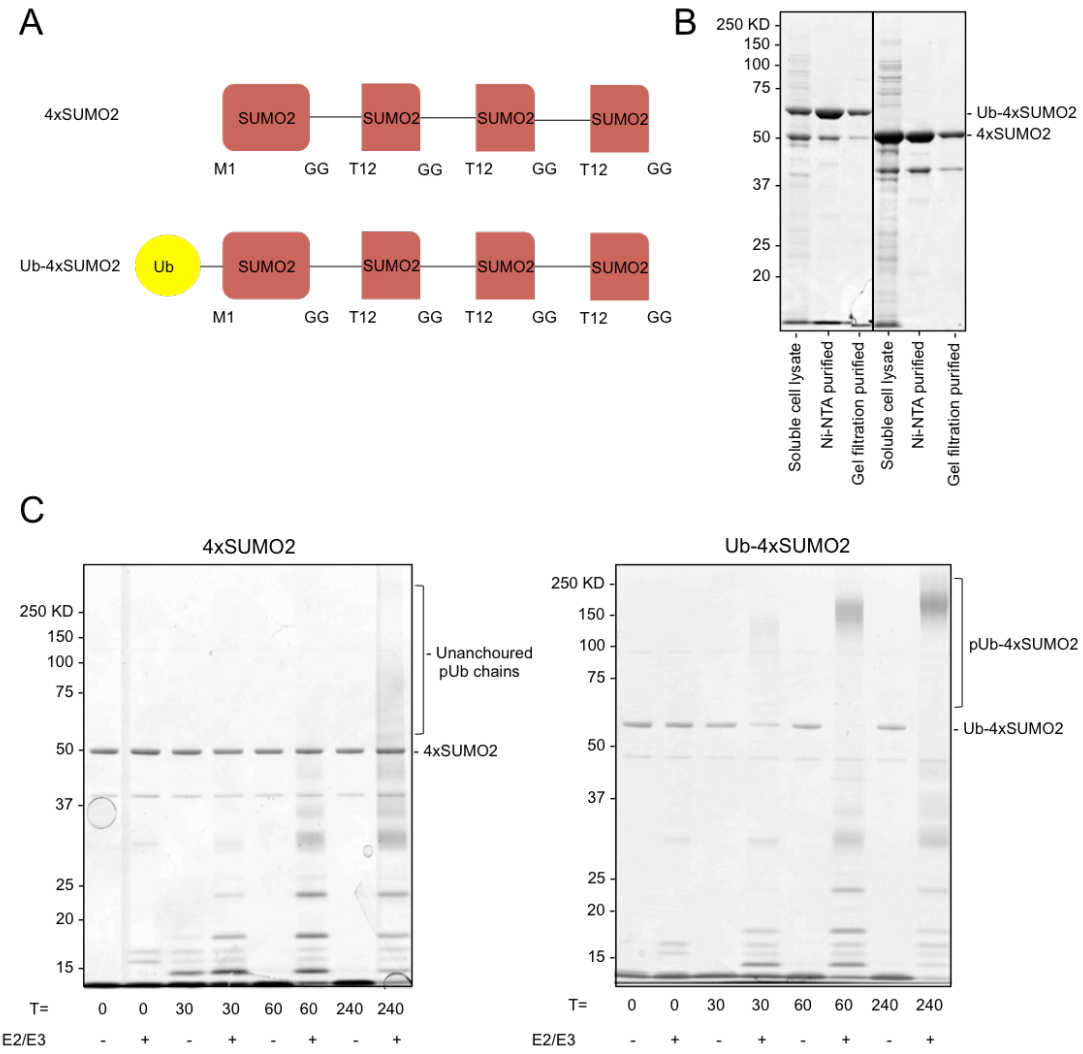


Figure 3.2-1 Generating hybrid chains of SUMO and ubiquitin.

A) Schematic diagram of linear 4xSUMO2 and Ub-4xSUMO2 constructs. Linear 4xSUMO2 construct contains 1 mature SUMO2 linked via peptide bonds sequentially to 3 modified SUMO2s with the first 11 N-terminal residues truncated. Ub-4xSUMO2 construct contains 1 ubiquitin moiety linked to the N-terminal of the mature full length SUMO of the linear 4xSUMO2 construct.

B) BL21 DE3 E.coli cells expressing either 4xSUMO2 or Ub-4xSUMO2 constructs were lysed. 4xSUMO2 and Ub-4xSUMO2 were then purified on Ni-NTA column before a further purification through a Superdex75 column (see **2.2.2.3**).

C) Coomassie stained SDS PAGE gel analysis of *in vitro* Ubiquitination of linear 4xSUMO2 and Ub-4xSUMO2. Ubiquitination reaction was performed using E2 enzymes Ubc13 and

UBE2v2 in conjunction with E3 enzyme RNF4. Reactions were carried out over the course of 4 hours at 37°C in the presence or absence of E2/E3 enzymes.

Unanchored K63 linked poly-ubiquitin chains, sourced from the Divison of Signal Transduction Therapy (DSTT, University of Dundee), consisting of chain lengths varying from ~3-10 ubiquitin moieties were used as a poly-ubiquitin chain type completing the required 'bait' proteins for our proteomic study. The four different chain types; poly-ubiquitin, poly-SUMO, mono-ubiquitin poly-SUMO, and poly-ubiquitin poly-SUMO (**Figure 3.2-2 A**) were then cross linked at equimolar concentrations to N-Hydroxysuccinimidyl activated agarose beads (NHS). This resulted in four different protein-containing resin types; K63 pUb, 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2. A protein free blank resin was also generated as a control against proteins interacting non-specifically with the NHS resins.

To test the efficacy of these resins as reagents capable of extracting proteins with affinities for SUMO and ubiquitin, recombinant proteins with known affinities for either SUMO or ubiquitin were used in pull-down assays to test each resin. RNF4 with its four N-terminally located SIM domains was used to test the affinity of the SUMO resins, while GST fusion protein containing two ubiquitin interacting motifs, GST-UIM dimer, was used to test the affinity for the ubiquitin portion of the resins. 5 µM of either RNF4 or GST-UIM dimer were incubated with 10 µl of each resin separately, any unbound proteins were washed from the resins before any protein interacting with the resins was eluted with SDS sample buffer. As expected, RNF4 was found to interact only with the three SUMO containing resin types; 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 but was not found to interact with the ubiquitin-only containing K63 pUb or the blank resin (**Figure 3.2-2 B, left**). GST-UIM dimer was found to interact only with the ubiquitin containing resin types; K-63 pUb, pUb-4xSUMO2, and to a lesser extent Ub-4xSUMO2 but did not interact with the SUMO-only containing 4xSUMO2 or blank resins (**Figure 3.2-2 B, right**). The results of these pull-down assays with recombinant SUMO or ubiquitin interacting proteins demonstrates that the K63 pUb, Ub-4xSUMO2

and pUb-4xSUMO2 bound resins are all capable of capturing proteins that show affinity for ubiquitin, while 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 are all capable of capturing proteins that show affinity for SUMO, as expected.

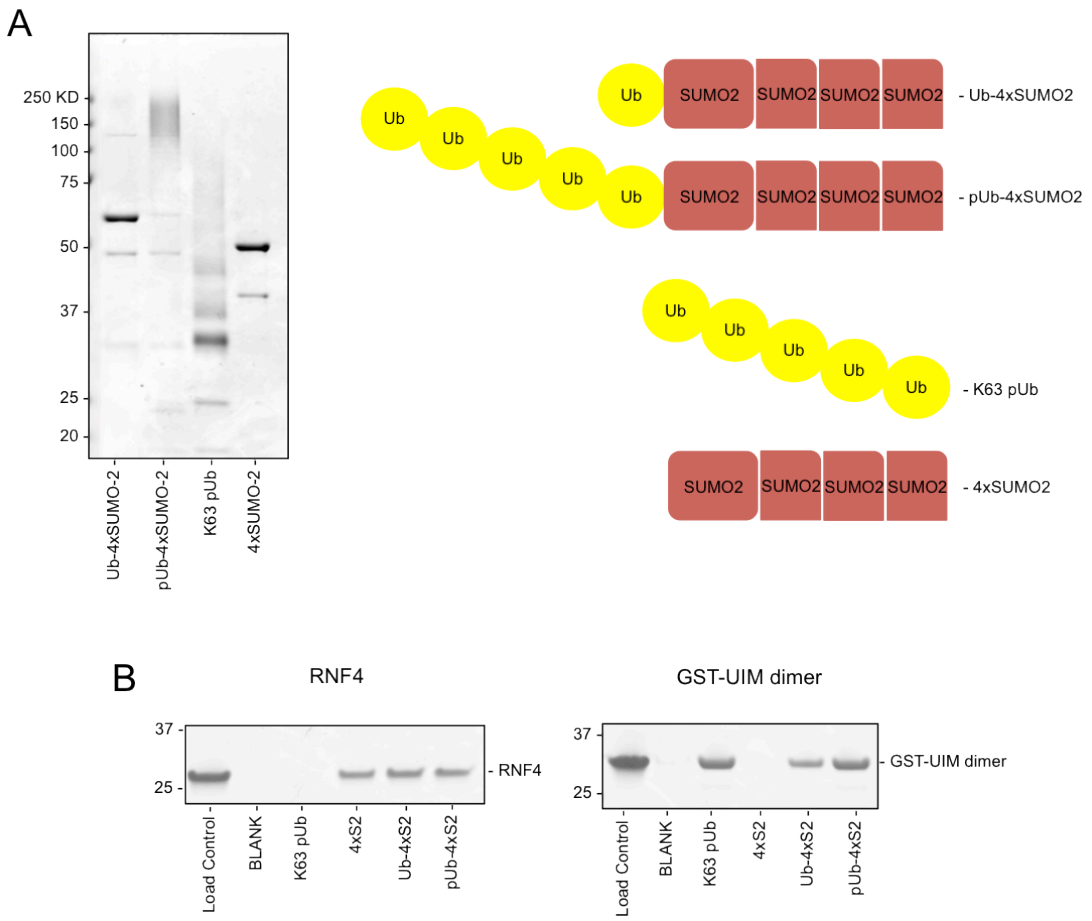


Figure 3.2-2 Interacting with 'bait' SUMO-ubiquitin hybrid chains.

A) Four different chain configurations were used to provide insight into SUMO, ubiquitin and SUMO-ubiquitin hybrid chain interacting proteins. Coomassie stained SDS PAGE gel analysis of the four different chain types; Ub-4xSUMO, pUb-4xSUMO2, K63 pUb, 4xSUMO2 prior to coupling to NHS-activated agarose resin. Schematic cartoon represents four different chain types used as bait in proteomic experiment; two hybrid chain types, Ub-4xSUMO2 and pUb-4xSUMO2; poly ubiquitin chain type, K63 pUb and poly SUMO chain type, 4xSUMO2.

B) Coomassie stained SDS PAGE gel analysis of pull down assays with SUMO specific interacting protein, RNF4, and ubiquitin specific interacting protein, GST-UIM dimer. 10µM recombinant RNF4 or GST-UIM dimer were incubated separately with 50µl of each resin type + blank resin before elution in 2xSDS sample buffer. RNF4 bound specifically to SUMO containing resins; 4xSUMO2, Ub4xSUMO2 and pUb-4xSUMO2. GST-UIM dimer interacts specifically with resins containing ubiquitin. Load control represent 20% of total protein incubated with resins.

3.2.2 A specific elution to elucidate information on interaction characteristic of hybrid chain interacting proteins

As a measure to try and elucidate more information about the affinity for SUMO or ubiquitin for any given potential interacting partner, a specific elution protocol was devised to dismantle SUMO-ubiquitin hybrid chains by sequentially removing the ubiquitin moieties before then removing the SUMO moieties of any chain. This would hopefully discern any solely ubiquitin interacting elements from any that were SUMO interacting, by first releasing proteins that were non-covalently linked to the ubiquitin portion of a hybrid chain before then releasing proteins the were non-covalently linked to the SUMO portion of the SUMO-ubiquitin hybrid chains (**Figure 3.2-3**).

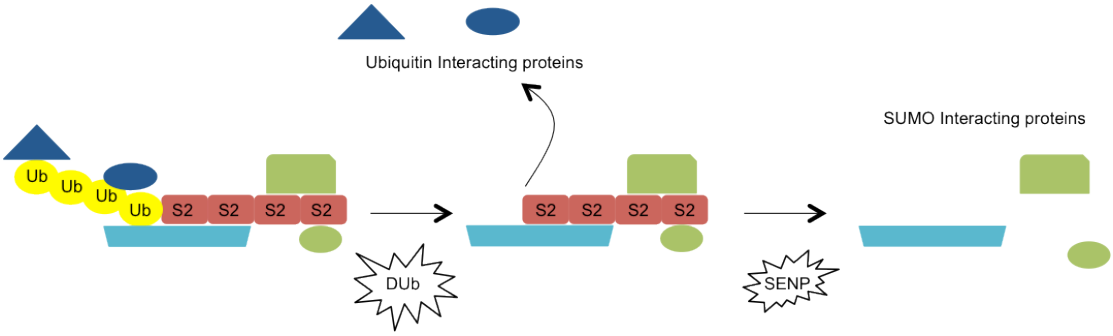


Figure 3.2-3 Stepwise elution of ubiquitin- then SUMO-interacting proteins by ubiquitin and SUMO proteases.

Cartoon represents sequential deubiquitnase and SUMO-specific protease based elution. Specific elution was designed to sequentially release ubiquitin-interacting proteins and then SUMO-interacting proteins from hybrid chains through a primary incubation with a highly active deubiquitinase enzyme followed by secondary incubation with a SUMO-specific protease enzyme.

Highly active, yet specific, SUMO-specific protease and deubiquitinase enzymes would need to be obtained to dismantle SUMO-ubiquitin hybrid chains in this manner. To this end, the active catalytic domain of the deubiquitinase enzyme, USP2, residues Asn258-Met605 (Referred to as USP2 CaD), was cloned and expressed in Rosetta ElaD (DE3) bacterial cells as described in **2.2.2.1**. Active SUMO-specific protease, SENP1 protein was a kind gift from Dr L. Shen (Shen et al., 2006).

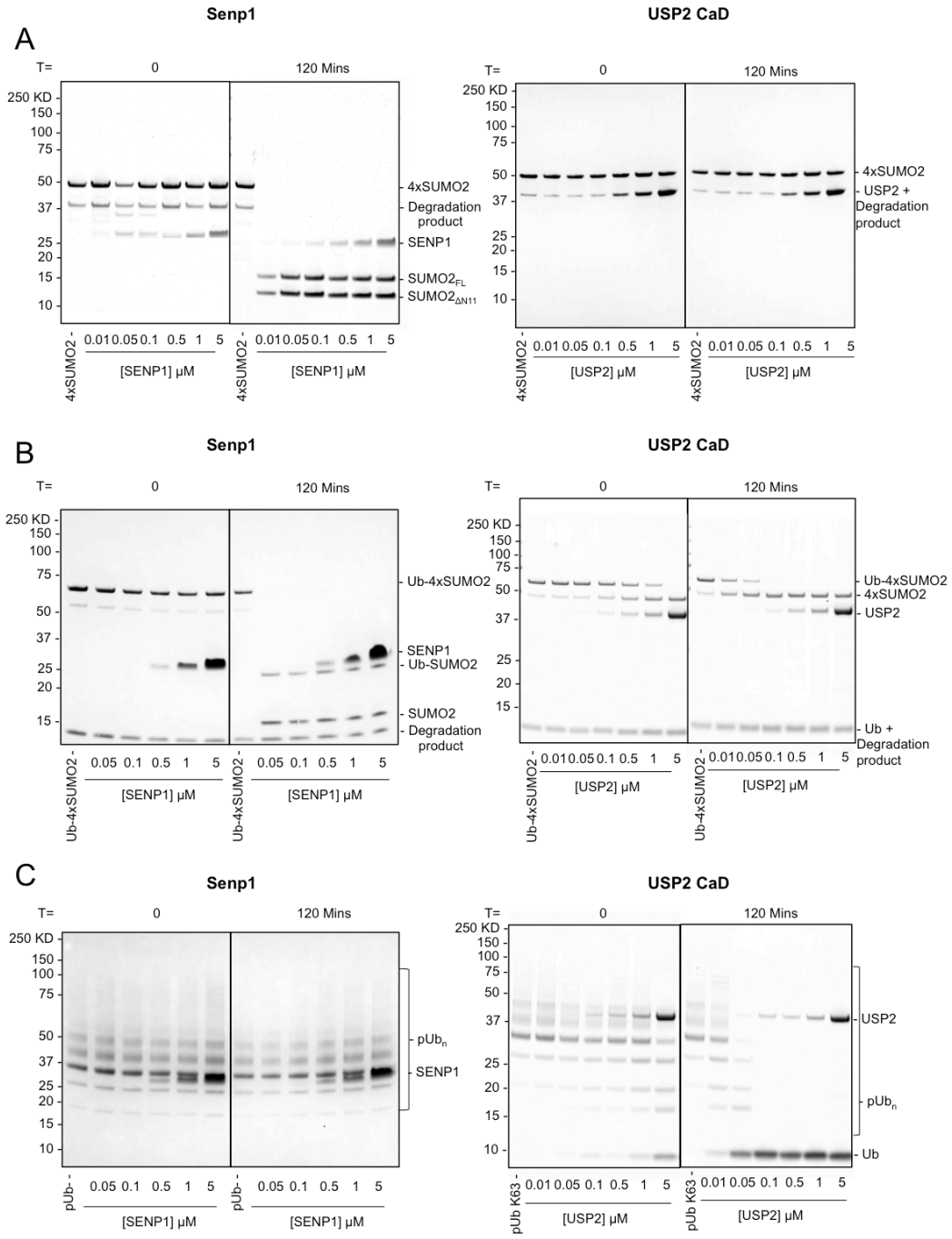


Figure 3.2-4 USP2 CaD and SENP1 are capabale of dismantling a SUMO-ubiquitin hybrid chain.

USP2 and SENP1 activity assays.

Stated concentrations of either SENP1 or USP2 were incubated with 10 µM;

A) 4xSUMO; B) Ub-4xSUMO2; C) K63 pUB for 0 or 120 minutes at 22° C. Substrate controls were incubated with no enzyme.

The activity of each enzyme was established against recombinant 4xSUMO2, Ub-4xSUMO2 and K63 pUb (**Figure 3.2-4**). As expected, after 120 minutes,

SENP1 had efficiently degraded the 4xSUMO2 poly-SUMO chain to mono-SUMO, at concentrations as low as 0.01 μM. After 120 minutes, USP2 CaD had had no effect on the 4xSUMO2 construct (**Figure 3.2-4, A**). As with the 4xSUMO2 poly-SUMO construct, SENP1 deconjugates SUMO from the Ub-4xSUMO2 construct. However, SENP1 does not show any activity towards ubiquitin-conjugated SUMO and so a secondary product corresponding to one ubiquitin covalently bound to one SUMO is also present after 120 minutes (**Figure 3.2-4, B left panel**). USP2 CaD at concentrations of 0.1 μM or above was capable of efficiently removing ubiquitin moiety from the Ub-4xSUMO2 construct, while not degrading the poly-SUMO element of the Ub-4xSUMO2 construct (**Figure 3.2-4, B right panel**). In line with expectations, SENP1 was unable to degrade K63 pUb chains *in vitro.* USP2 CaD, however, efficiently degrades the K63 pUb chains to mono-ubiquitin in 120 minutes at concentrations of 0.1 μM or above (**Figure 3.2-4, C**). Thus, the purified USP2 CaD and SENP1 enzymes provide the catalytic activity capable of degrading SUMO-ubiquitin hybrid chains efficiently in a stepwise and controlled manner.

The stepwise USP2 CaD/SENP1 elution protocol was then tested against the K63 pUb, 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 containing resins incubated with the SUMO- and ubiquitin-interacting proteins, RNF4 and GST-UIM dimer. Resins were incubated with both RNF4 and GST-UIM dimer overnight prior to elution. The three-stage elution process was then carried out; USP2 CaD was incubated with each resin for 2 hours. The resins were spun down and the USP2 CaD supernatant, containing any proteins that were dislodged by the protease treatment were removed. SENP1 was then incubated with each resin for 2 hours, again the resins were spun down and the SENP1 supernatant and any proteins dislodged by the treatment were removed. Finally the resins were boiled in SDS sample buffer to release any proteins still interacting with the resins, before all elution steps were analyzed by SDS-PAGE and western blotting separately (**Figure 3.2-5**)

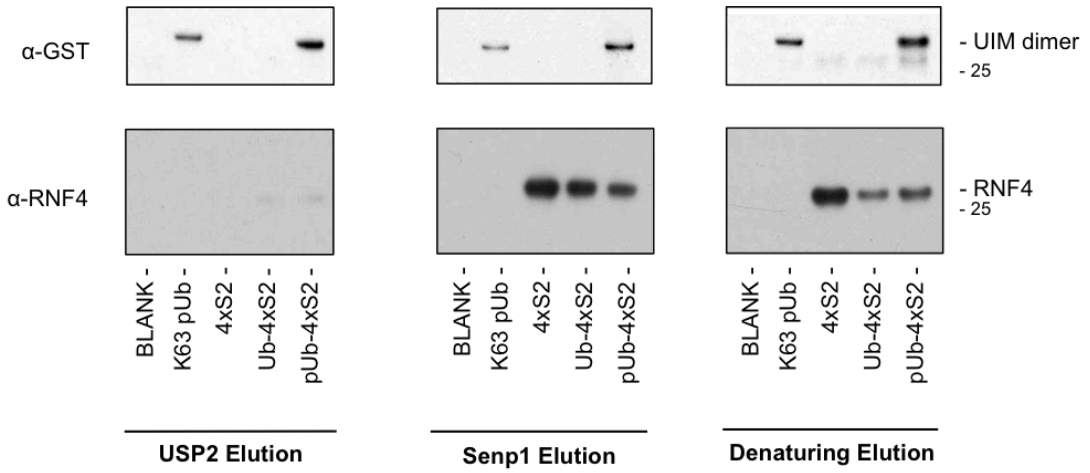


Figure 3.2-5 USP2/SENP1 Specific elution from resins incubated with SUMO or ubiquitin specific interacting proteins.

5 μM RNF4 and GST-UIM dimer were incubated together with 50 μl of each of the 4 resin types + Blank resin separately, overnight at 4°C. Step wise elution with 5 μM USP2 followed by 5 μM SENP1 and a final SDS boil was carried out for each resin and analysed by western blot via indicated antibody.

As expected, GST-UIM dimer was released from the K63 pUb and pUb-4xSUMO2 resins after treatment with USP2 CaD. No GST-UIM dimer was recovered from the 4xSUMO2 resin or, surprisingly, from the Ub-4xSUMO2 resin. The addition of SENP1 to the resins released a sizeable amount of RNF4 from the 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 resin types, but consistent with the previous pull downs, no RNF4 was eluted from the K63 pUb resin types. GST-UIM dimer was also present in the SENP1 elution from the K63 pUb and pUb-4xSUMO2 resins but this could be accounted for by residual GST-UIM dimer left over from the USP2 CaD elution still being present. RNF4 was found in the eluate of the 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 resins after the resins had been boiled in SDS, with GST-UIM dimer present in the SDS elution from K63 pUb and pUb-4xSUMO2 resin types.

Thus, an elution based around disrupting SUMO ubiquitin hybrid chains via the protease activity of USP2 CaD and SENP1 was capable of dislodging proteins that were known to be non-covalently interacting specifically with

either the ubiquitin moiety or the SUMO moiety of a SUMO-ubiquitin hybrid chain.

The identification of hybrid chain interacting proteins from complex cell lysates was then key to understanding whether or not the combination of SUMO-ubiquitin hybrid chain pull down assay and USP2 CaD/SENP1 elution protocol was capable of identifying novel hybrid chain interacting proteins. As such the SUMO-ubiquitin hybrid chain pull down assay and USP2 CaD/SENP1 elution protocol were tested after incubation with nuclear cell lysates.

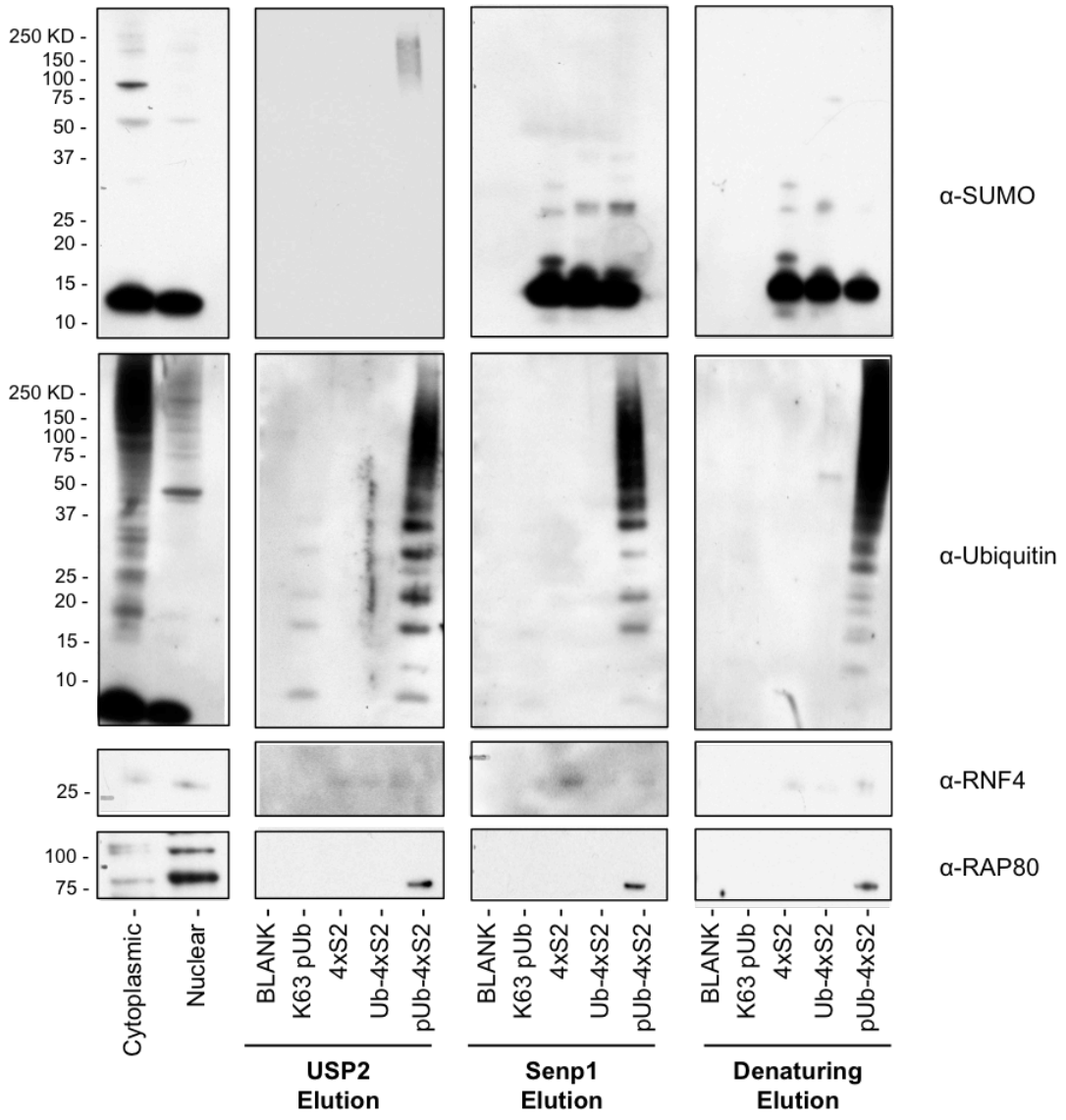


Figure 3.2-6 Hybrid chain pull down with HEK 293 N3S nuclear lysates.

HEK 293 N3S nuclear extracts were incubated separately with BLANK, K63 pUb, 4xSUMO-2, Ub-4xSUMO-2 and pUb-4xSUMO-2 resins. USP2 CAD, SENP1 and SDS elutions were carried out sequentially for each resin before being analyzed by Western blot for indicated antibody. Pre- incubation nuclear extract and a cytoplasmic extract act as loading controls.

HEK 293 spinner cell cultures (5 litres) were grown to a density of 1E6 cells/ml before the cells were spun down and lysed into cytoplasmic and nuclear extracts in accordance with **2.2.5.2**. Each of the five resin types were then incubated overnight with HEK 293 nuclear cell extracts before the USP2 CaD/SENP1 elution protocol was performed. The elutions were then separated by SDS-PAGE and analysed by western blot (**Figure 3.2-6**) with antibodies against ubiquitin, SUMO, RNF4, and the previously identified potential SUMO-ubiquitin hybrid chain interacting protein RAP80 (Guzzo et al., 2012). Free monomeric SUMO was found to be present after the incubation with SENP1 and the denaturing elution of the SUMO containing resin types 4xSUMO2, Ub-4xSUMO2, and pUb-4xSUMO2. Ubiquitin was found in the eluate from the ubiquitin containing resin types, K63 pUb and pUb-4xSUMO2 after incubation with USP2 CaD, SENP1, and the denaturing elution. In line with pervious experiments RNF4 was found to be present in the eluate from 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 after the SENP1 elution and the denaturing elution, with the most prominent RNF4 band detected with the 4xSUMO2 resin after SENP1 elution. RNF4 was not found to be present in any elutate from blank or K63 pUb resin types. Interestingly, RAP80 was detected in the poly-SUMO poly-ubiquitin hybrid chain containing pUb-4xSUMO2 resin type after all three elutions but not in any elution from the Ub-4xSUMO2 hybrid chain type, 4xSUMO2, K63 Ub, or blank resins. This suggests that potential hybrid chain interacting proteins may only interact with specific forms of SUMO-ubiquitin hybrid chains.

Identifying RAP80 by western blot after SUMO-ubiquitin hybrid chain pull-down assay gives credence to the hypothesis that a combination of the 4xSUMO2, K63 pUb, and Ub-4xSUMO2 and pUb-4xSUMO2 hybrid chain resin types along with the specific USP2 CaD/SENP1 elution protocol is

capable of identifying novel SUMO-ubiquitin hybrid chain interacting proteins from complex human cell lysates.

3.2.3 Detection of SUMO-ubiquitin hybrid chain binding proteins via protein affinity chromatograph coupled with high resolution mass spectrometry

Having identified RAP80 by western blot analysis in the elution's from HEK 293 cells lysates using the SUMO-ubiquitin hybrid chain pull-down assay outlined above, samples were then prepared for mass spectrometry analysis in the hope of identifying a number of novel potential SUMO-ubiquitin hybrid chain interacting proteins.

In total 45 litres of HEK 293 spinner cells were cultured before nuclear extracts were prepared. 150 µl of each resin type was incubated with 10 mL nuclear lysates (5.5 mg/ml) with rotation overnight at 4°C. Resins were then washed to remove any proteins not specifically bound before incubation with USP2 CaD, SENP1 and a denaturing elution with LDS sample buffer were used to elute any bound proteins from the resins. The eluates from each resin for each elution, 15 in total, were then separated by SDS-PAGE (**Figure 3.2-7 A**). After coomassie staining the elution profiles of resin were compared. Noticeable differences were apparent in the elution profiles from each resin type, with the majority of the material being released from the resins by the denaturing LDS sample buffer boil. The lanes containing each elution were then cut into two slices and prepared for mass spectrometry analysis by in gel trypsin digestion (**2.2.8.1**). Samples were then processed using MaxQuant, reporting intensities for 2197 proteins. The intensities reported by MaxQuant can be used as a surrogate for the relative abundance for each protein, from each elution, for each resin type. This can then be used to describe the apparent relative affinity that each protein shows for each resin type i.e. if after all elutions are carried out a protein is reported to have a large intensity in the pUb-4xSUMO2, Ub-4xSUMO2, and 4xSUMO2 resin types but not in BLANK or K63 pUb resin types then that protein can be said to show affinity for the 4xSUMO2 containing resin types. To remove any bias from the analysis of any potential SUMO-ubiquitin hybrid chain

interacting protein numerical analysis was undertaken on the mass spectrometry data to determine if any proteins identified in the pull-down assay could be considered to be interacting specifically with poly-SUMO, poly-ubiquitin, or with either form of SUMO-ubiquitin hybrid chain. This was achieved by scoring each protein based on which elution, and from which resin type that protein was identified in Table 3.1.1 illustrates the scoring system. Firstly, any protein with a missing intensity score was replaced with 1E6, the lower limit for intensity report. For each experiment all 10 ratio combinations of intensity scores for; blank (A), K63 pUb (B), 4xSUMO2 (C), Ub-4xSUMO2 (D), and pUb-4xSUMO2 (E) were calculated, and the data normalized from frequency histograms of each $\log_2$ ratio. Using this scheme a 'specificity score' was calculated for each protein for each resin type. This gives a score for the apparent consistency of the intensity data with that expected of a protein binding to the resins with specific characteristics, such as 'binds to any poly-SUMO resin'. In theory these scores are comparable, with the maximum score indicating the predicted specificity of that protein. For each protein the maximum specificity score was calculated for each resin. These maxima were used to create frequency histograms to see the spread of maximum values for each experiment. These data were used to plot normal distributions and create 95% confidence intervals. So all maximum specificity scores >~6 were regarded as significant. Example specificity score for 'Any polySUMO'

'$=((\sqrt{(\log_2 B/A)^2+\log_2 C/A+\log_2 D/A+\log_2 E/A+\log_2 C/B+\log_2 D/B+\log_2 E/B+\log_2 D/C+\sqrt{(\log_2 E/C)^2+\sqrt{(\log_2 E/D)^2})*\log_{10}}$Total Protein Intensity-6'. This score takes into account all ratio data as well as the abundance of the protein (Total protein intensity). -6 allows for the fact that the minimum protein intensity is 1E6.

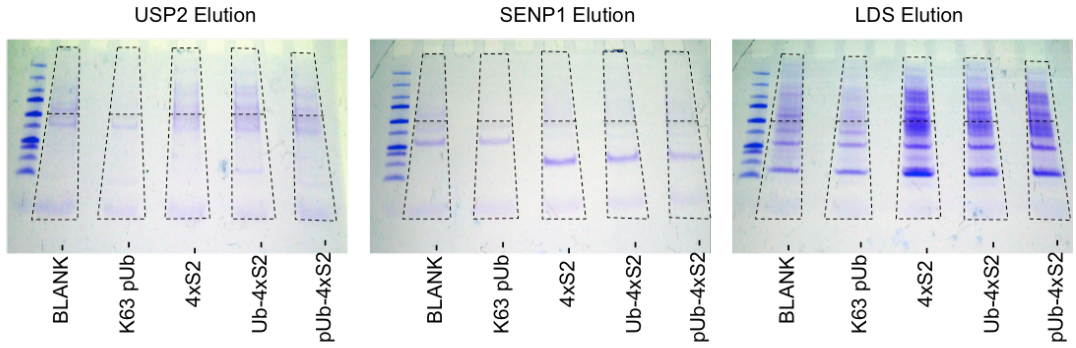
Specificity	B/A	C/A	D/A	E/A	C/B	D/B	E/B	D/C	E/C	E/D	# Contributing ratios (n)
Any polySUMO	0	+	+	+	+	+	+	+	0	0	6
Any Ub-PolySUMO	0	0	+	+	0	+	+	+	+	0	6
Ub-polySUMO	0	0	+	0	0	+	0	+	0	-	4
PolyUb-PolySUMO	0	0	0	+	0	0	+	0	+	+	4
Any K63 Ub	+	0	0	+	-	-	0	0	+	+	6
Any Ub	+	0	+	+	-	0	0	+	+	0	6

Table 3.2-1 – Specificity scoring system for protein affinities.

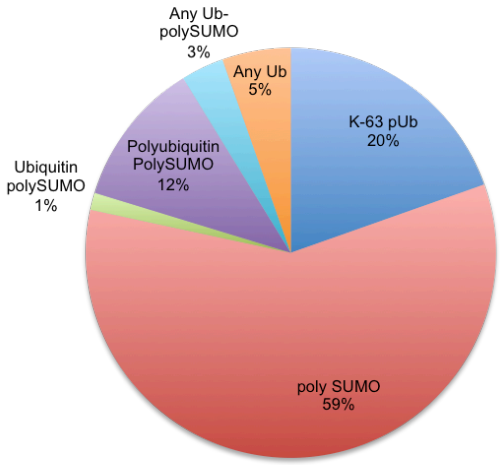
Table outlines the predicted ratios expected for a given interaction characteristic. (0) Describes a ratio that would not be expected to contribute towards a certain interaction characteristic. (–) Describes a ratio that would be expected to be negative for the described interaction characteristic. (+) Describes a ratio that would be expected to be positive for the described interaction characteristic. Example – specificity score for 'Any polySUMO' '=((√(log_2B/A)^2+log_2C/A+log_2D/A+log_2E/A+log_2C/B+log_2D/B+log_2E/B+log_2D/C+√(log_2E/C)^2+ √(log_2E/D)^2)*log_{10}Total Protein Intensity-6'. This score takes into account all ratio data as well as the abundance of the protein (Total protein intensity). -6 allows for the fact that the minimum protein intensity is 1E6.

This method resulted in a list of 152 proteins that could be described as having defined interaction specificity. For each of these the highest maximum of the three elutions from each resin was considered as the predominant specificity and reported. 59% of the proteins identified as having an affinity for one of the resin types were characterised as any poly-SUMO interacting proteins (**Appendix II**). 25% of the proteins were identified as ubiquitin interacting proteins, 5% were interacting with any ubiquitin moiety, with the remaining 20% listed as K63 specific i.e. interacting with the poly-ubiquitin containing resin types K63 pUb and pUb-4xSUMO2 only (**Appendix II**). In total, 16% of the total proteins listed as having specificity were suggested to be interacting with a form of SUMO-ubiquitin hybrid chain (**Figure 3.2-7, B**). 12% of the total number of proteins were said to be interacting specifically with the poly-SUMO poly-ubiquitin hybrid chain type. Only 1% was suggested to show specificity for the mono-ubiquitin poly-SUMO hybrid chain type, while the remaining 3% had affinity for both the mono- and poly-ubiquitin containing hybrid chain types. Along with the numerical analysis of the data the intensity charts for each protein identified by MaxQuant were also checked manually, by comparison with that of the known potential SUMO-ubiquitin hybrid chain interacting protein RAP80 (**Figure 3.2-8, A**). Any protein that was not identified by the numerical analysis but had a similar intensity chart to that of RAP80 was also added to the list of potential SUMO-ubiquitin hybrid chain interacting proteins (denoted by *), bringing the number of potential SUMO-ubiquitin hybrid chain interacting proteins to 30 (**Figure 3.2-7, C**).

A



B



C

Protein names	Gene names	Identifer
Tubby-related protein 4	TULP4	UbpS
Zinc finger MYM-type protein 2	ZMYM2	UbpS
Activating signal cointegrator 1 complex subunit 3	ASCC3	pUbpS
BRCA1-A complex subunit Abraxas	FAM175A	pUbpS
BRCA1-A complex subunit BRE	BRE	pUbpS
BRCA1-A complex subunit RAP80	RAP80/UIMC1	pUbpS
BRISC and BRCA1-A complex member 1	BABAM1	pUbpS
Breast cancer type 1 susceptibility protein	BRCA1	pUbpS
Lys-63-specific deubiquitinase BRCC36	BRCC3	pUbpS
Next to BRCA1 gene 1 protein*	NBR1*	pUbpS
Centromere-associated protein E	CENPE	pUbpS
Coiled-coil domain-containing protein 50	CCDC50	pUbpS
E3 ubiquitin-protein ligase HUWE1	HUWE1	pUbpS
E3 ubiquitin-protein ligase RNF168	RNF168	pUbpS
Epidermal growth factor receptor substrate 15-like 1	EPS15L1	pUbpS
Epsin-1	EPN1	pUbpS
Epsin-2	EPN2	pUbpS
MORC family CW-type zinc finger protein 3*	MORC3*	pUbpS
MORC family CW-type zinc finger protein 4	MORC4	pUbpS
Sequestosome-1	SQSTM1	pUbpS
Tubulin beta-6 chain	TUBB6	pUbpS
Zinc finger protein 451	ZNF451	pUbpS
Protein ELYS	AHCTF1	AUbpS
Complexin-3	CPLX3	AUbpS
Integrator complex subunit 7	INTS7	AUbpS
Obscurin	OBSCN	AUbpS
Uncharacterized protein C5orf25	C5orf25	AUbpS
E3 ubiquitin-protein ligase RNF31*	RNF31*	pUbpS
E3 ubiquitin-protein ligase TRIM32*	TRIM32*	pUbpS

Figure 3.2-7 Mass spectrometry analysis of hybrid chain interacting proteins.

A) SDS-PAGE analysis of eluate from resins types after incubation with 293 N3S nuclear extracts. The three step specific elution was carried out and eluate from each resin type were run on 10% SDS-PAGE gel before individual lanes were sliced as shown in black dashed boxes.

B) Pie chart summation of specific SUMO, ubiquitin, or SUMO ubiquitin hybrid chain interacting proteins identified by mass spectrometry. PolySUMO describes any protein that showed affinity only for the 4xSUMO2 resin. K63 pUb describes any protein that showed affinity only for the K63 pUb resin. Any Ub describes proteins that show affinity for any of the ubiquitin containing resins. Any Ub-polySUMO describes proteins that showed affinity for both the Ub-4xSUMO2 and pUb-4xSUMO resins. Polyubiquitin polySUMO describes any protein

that showed affinity only for the pUb-4xSUMO2. Ubiquitin polySUMO describes any protein that showed affinity only for ub-4xSUMO2 resin.

C) Table listing all proteins that showed affinity for a hybrid chain type resin. Proteins highlighted with (*) were added to the initial list because of similarity of their intensity charts to that of RAP80. Proteins taken for further investigation are highlighted in red, with the BRCA 1A complex members are highlighted in blue.

Along with RAP80, five other BRCA 1A complex components were characterized as SUMO-ubiquitin hybrid chain interacting proteins, including Abraxas, BRE, BRCC36, BABAM1 (Merit40) and BRCA 1 (**Figure 3.2-7, C, highlighted in blue**). BRE, Abraxas, and BRCC36 are known to contain ubiquitin interacting elements. BRE contains two ubiquitin-conjugating enzyme like (UEV) domains (Wang et al., 2009). Abraxas and BRCC36 contain JAMM/MPN ubiquitin binding domain (MPN) like domains (Sobhian et al., 2007). The sequence of each of these proteins was searched for SIM domains using the PATTINPROT web tool. Potential SIMs were found in BRE, Abraxas, and BRCC36 (**Figure 3.2-8, B, dashed red lines**). This suggests that BRE, Abraxas and BRCC36 have the potential, like RAP80, to interact with both SUMO and ubiquitin. The combined intensity charts, which report the combined intensities for all elutions for each resin type minus any reported intensities from the blank resin, for the BRCA 1A complex components all display a similar pattern of affinity for the pUb-4xSUMO2 resin (**Figure 3.2-8, C**). The intensity charts for each of these proteins would suggest that they each show affinity for the same poly-SUMO poly-ubiquitin form of SUMO-ubiquitin hybrid chain. This may suggest that this type of SUMO-ubiquitin hybrid chain is very important to the role of the BRCA 1A complex.

The identification of RAP80 in the mass spectrometry data as a SUMO-ubiquitin hybrid chain interacting protein, reinforces the hypothesis that a combination of the 4xSUMO2, K63 pUb, and Ub-4xSUMO2 and pUb-4xSUMO2 hybrid chain resin types along with the USP2 CaD/SENP1 elution protocol is capable of identifying novel hybrid chain interacting proteins from complex human cell lysates.

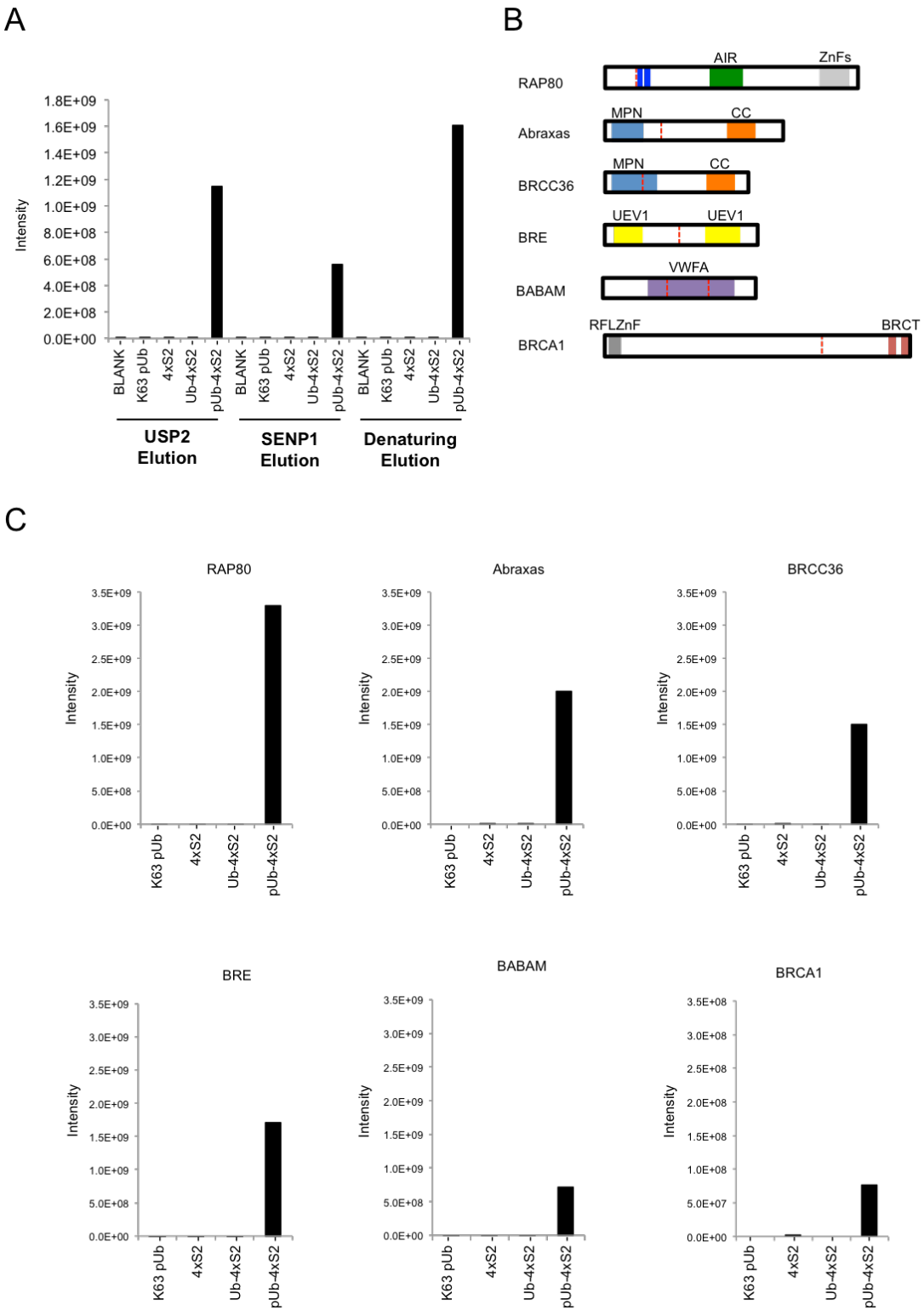


Figure 3.2-8 Proteomic analysis of BRAC 1A complex components.

A) MaxQuant reported intensity charts for RAP80 from each of USP2, SENP1 and denaturing elutions. Resin type is listed below graph. Elution type is listed below resin type.

B) Schematic depiction of BRCA1 A complex components RAP80, Abraxas, BRCC36, BRE, BABAM/Merit40 and BRCA1. Potential or described SUMO interaction motifs are shown with

dashed red line. Ubiquitin interaction motifs are shown by dark blue boxes. Green box outlines AIR synthase domain of RAP80. Grey box indicates Zinc fingers of RAP80 and BRCA1. Blue boxes indicate MPN domains of Abraxas and BRCC36. Orange boxes indicate coiled coil domains. Yellow boxes indicate relative position of UEV1 domains with BRE. Purple box of BABAM/Merit40 represents VWFA domain. Red boxes of BRCA1 represent BRCT domains.

C) MaxQuant reported intensity charts for each for the BRCA 1A complex components. Reported intensities for each elution combined together for each protein containing resin minus the combined intensities from USP2, SENP1 and denaturing elutions from BLANK resin in each case.

3.2.4 Verifying potential hybrid chain interacting proteins

The remaining proteins identified as showing affinity for SUMO-ubiquitin hybrid chains were then systematically appraised for the domains required to interact with both SUMO and ubiquitin. Literature and database searches were carried out for any known ubiquitin binding domains or proteins known to contain SIMs. This revealed that a number of the proteins in the hybrid chain interacting list were already described to contain either SUMO- or ubiquitin-interacting elements. Three of these proteins were selected for further analysis C5orf25/SIMC1, RNF168 and MORC3 (**Figure 3.2-7, C, highlighted in red**).

C5orf25/SIMC1 (referred to as C5orf25) is known to contain two high fidelity type SIM domains located in its N-terminal region but has no known ubiquitin binding domains (**Figure 3.2-9, A**) (Sun & Hunter 2012). The combined intensity chart for C5orf25 suggests that the protein shows affinity for the hybrid chain resin types, Ub-4xSUMO2 and pUb-4xSUMO2, but also for the poly-SUMO only containing resin 4xSUMO2 (**Figure 3.2-9, B**). Interestingly, when analysing the intensity chart from the USP2 CaD, SENP1 and denaturing elution's separately (**Figure 3.2-9, C**), intensities are reported for C5orf25 in USP2 CaD elution of the hybrid chain resin types pUb-4xSUMO2 and Ub-4xSUMO2. This suggests that the ubiquitin moieties of those resin types have some affect on the ability of C5orf25 to interact with those resins. No intensity was reported for C5orf25 in the USP2 CaD elution

from the poly-SUMO only 4xSUMO2 resin type. This suggest the C5orf25 released from the hybrid chain resins after USP2 CaD elution is not due to an unspecific effect of USP2 acting on C5orf25. If this were the case, C5orf25 would be released from all the resin types by the USP2 CaD elution. Intensities are reported for C5orf25 from the 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 resins after the SENP1 elution. C5orf25 released by this elution from the 4xSUMO2 resin type must be due to the disruption of the SUMO chains of this resin by the SENP1 enzyme suggesting that as expected, C5orf25 has the ability to interact with SUMO, most likely via the two previously identified high fidelity SIM domains. As expected, intensities are also reported for C5orf25 from the 4xSUMO2, Ub-4xSUMO2 and pUb-4xSUMO2 resins after the LDS sample buffer elution. The highest intensity reported for C5orf25 were recorded in eluates from the pUb-4xSUMO2 resin type, suggesting C5orf25 was most abundant in the eluate from that resin type. Thus, C5orf25 showed the greatest affinity for the poly-SUMO poly-ubiquitin hybrid chain resin pUb-4xSUMO2.

These data suggest that C5orf25 can interact with poly-SUMO, but not with K63 linked poly-ubiquitin alone. C5orf25 however, does appear to have the greatest affinity for hybrid chains of SUMO and ubiquitin, with the ubiquitin moieties of these chains playing some role in the interaction. With this in mind, recombinant [^{35}S] methionine labelled C5orf25 was generated via an *in vitro* transcription/translation reaction. The recombinant C5orf25 was then tested in a SUMO-ubiquitin hybrid chain pull-down assay (**Figure 3.2-9, D**). Recombinant C5orf25 interacts in a similar manner to the C5orf25 from the HEK 293 cell extracts, showing affinity for all the SUMO containing resin types. A slight increase in C5orf25 was observed in the SUMO-ubiquitin hybrid chain resin types when compared to the 4xSUMO2 resin type. C5orf25 does appear to be a bona fide SUMO-ubiquitin hybrid chain interacting protein. However, due to lack of any apparent ubiquitin binding domain further study would be required to identify any potential domains that could facilitate this interaction.

Another hit from the hybrid chain interacting screen was the ubiquitin E3 ligase, RNF168. RNF168 is known to have a role in the DNA damage response, potentially upstream of the recruitment of RAP80, making it an interesting target for further study (Doil et al., 2009). RNF168 is known to have three UIMs, one canonical UIM and two reverse UIMs, known as MIUs (Pinato et al 2009, Panier et al., 2012). Previous work in our lab carried out by Dr. J-F Maure (unpublished data) had suggested that RNF168 has four potential SIM domains, two of these potential SIM domains and are located in close proximity to the secondary MIU of RNF168 (**Figure 3.2-9, E**), which is required for the recruitment of RNF168 to sites of DNA repair (Doil et al 2009). This set up of two SIM domains and one MIU could act in a similar way to the tandem SIM UIMUIM domain of RAP80 and facilitate a specific interaction between RNF168 and SUMO-ubiquitin hybrid chains. The combined elution profile of RNF168 suggests that it has affinity for the poly-SUMO poly-ubiquitin hybrid chain containing, pUb-4xSUMO2 resin (**Figure 3.2-9, F**). Intensities were reported for RNF168 in the elutions from the 4xSUMO2 resin type but no RNF168 was detected in either the K63 pUb or Ub-4xSUMO2 resin types. Only a small amount of RNF168 was detected in the eluate from the USP2 CaD elution from the pUb-4XSUMO2 resin type. The largest intensities for RNF168 were reported in the eluate from the denaturing elution from the pUB-4xSUMO2 resin, with smaller intensities reported for RNF168 in the eluate of the 4xSUMO2 resin after denaturing elution (**Figure 3.2-9, G**). Bacterially expressed recombinant RNF168 was then tested in the SUMO-ubiquitin hybrid chain pull-down assay (**Figure 3.2-9, H**). The most prominent RNF168 band was associated with the eluate of the pUb-4xSUMO2 resin, agreeing with the data from the mass spectrometry analysis. Unlike the mass spectrometry data the second most prominent RNF168 band was associated with the elution from the K63 pUb resin type. This however agrees with current literature that suggests that RNF168 can interact with K63 linked poly-ubiquitin chains (Panier et al., 2012).

RNF168 appears to show affinity for poly-SUMO poly-ubiquitin hybrid chain types developed for this assay. This is perhaps not surprising given

that RNF168 contains three known ubiquitin interacting motifs and a number of potential SIM domains.

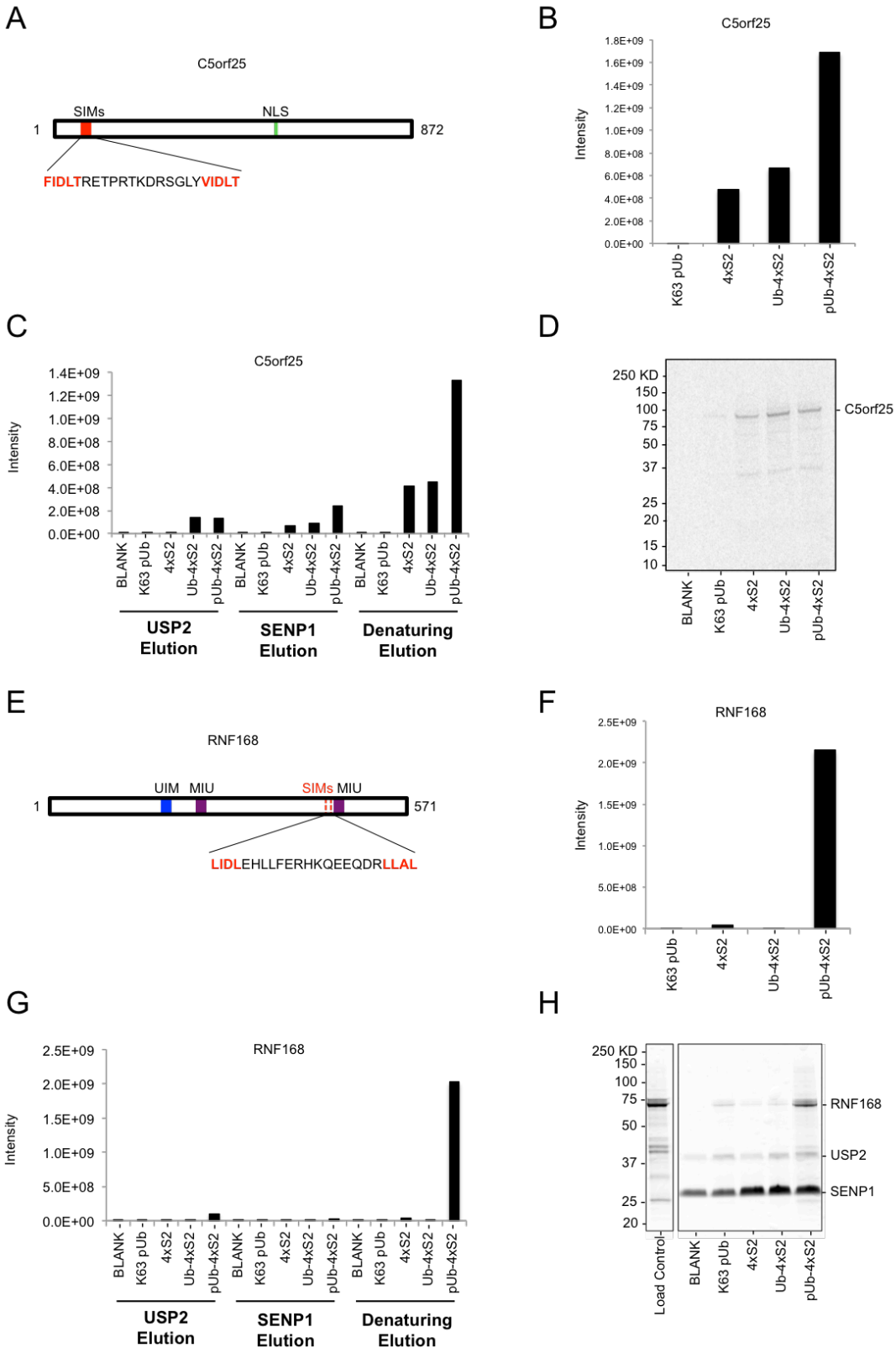


Figure 3.2-9 Exploring novel potential hybrid chain interacting proteins.

A) Schematic diagram represents C5orf25. Two N-terminally located SUMO interaction motifs are represented by red box, sequence is shown highlighted in red below. Nuclear localisation signal is shown by green line.

B) Combined intensities reported by MaxQuant for C5orf25 minus intensities reported for blank resin.
C) Reported intensities from MaxQuant for C5orf25 for each resin type from each elution.
D) Hybrid chain interacting protein pull-down assay performed with recombinant C5orf25 produced by in vitro transcription and translation
E) Schematic diagram represents RNF168. Ubiquitin interaction motif represented by blue box with reverse ubiquitin interaction motifs shown by purple boxes. Potential SIMs shown by dashed red lines, sequence highlighted in red below.
F) Combined intensities reported by MaxQuant for RNF168 minus intensities reported for blank resin.
G) Reported intensities from MaxQuant for RNF168 for each resin type from each elution.
H) Hybrid chain interacting protein pull-down assay performed with recombinant RNF168. Coomassie stained SDS PAGE gel analysis of hybrid chain pull-down assay with RNF168. 10μM recombinant RNF168 were incubated with 50μl of each resin type + blank resin. USP2 and SENP1 elution were carried out before a final elution in 2xSDS sample buffer. Load control represents 20% of total recombinant protein incubated with beads

---

Two members of the MORC family CW-type zinc finger proteins, MORC3 and MORC4, were identified as potential hybrid chain interacting proteins. Of these two, MORC3 has previously been described to be associated with PML in PML-NBs (Mimura et al., 2010). PML is known to be SUMOylated and to be acted on by the STUbL RNF4 (Duprez et al.,1999, Tatham et al., 2008) and therefore has the potential to anchor hybrid chains of SUMO and ubiquitin. The apparent affinity of MORC3 for SUMO-ubiquitin hybrid chains was therefore very interesting and worth further investigation.

MORC3 is known to be SUMOylated at two positions; K597 and K650 (Tammsalu et al., 2014). 9 potential SIMs are found throughout the sequence of MORC3 (**Figure 3.2-10, A**). However, MORC3 contains no known ubiquitin binding domains. The combined elution profile of MORC3 is therefore quite surprising, with the largest signal intensity is reported from the pUb-4xSUMO2 resin type, and with smaller intensities being reported from the elutions from both the 4xSUMO2 and Ub-4xSUMO2 resins (**Figure 3.2-10, B**). This suggests that although MORC3 shows some affinity for poly-SUMO, it shows a far greater affinity for poly-SUMO poly-ubiquitin hybrid chains. To

test how reproducible these finding were *in vitro*, bacterially expressed MORC3 was tested in the SUMO-ubiquitin hybrid chain pull-down assay (**Figure 3.2-10, C**). The apparent affinity of the recombinant MORC3 for each resin type was remarkably similar to that shown by the elution profile in the combined intensity chart reported for MORC3 (**Comparing Figure 3.2-10, B to D**). MORC3 was significantly more abundant in the elution from the pUb-4xSUMO2 resin type when compared to the Ub-4xSUMO2 resin, 4xSUMO2 resin, or K63 pUb resin type (**Figure 3.2-10, D**), with MORC3 found to be more abundant in the elutions from the SUMO containing resins and showing no real affinity for the K63 pUb ubiquitin only resin type. This agrees with the current literature that suggests MORC3 is a SUMO interacting protein (Mimura et al., 2010). The lack of affinity towards the pUb K63 ubiquitin chains is not surprising given the lack of a defined ubiquitin binding domain in MORC3, however this only adds to the intrigue as to how the apparent affinity of MORC3 for hybrid chains of SUMO and ubiquitin is so much more intense than for SUMO chains alone.

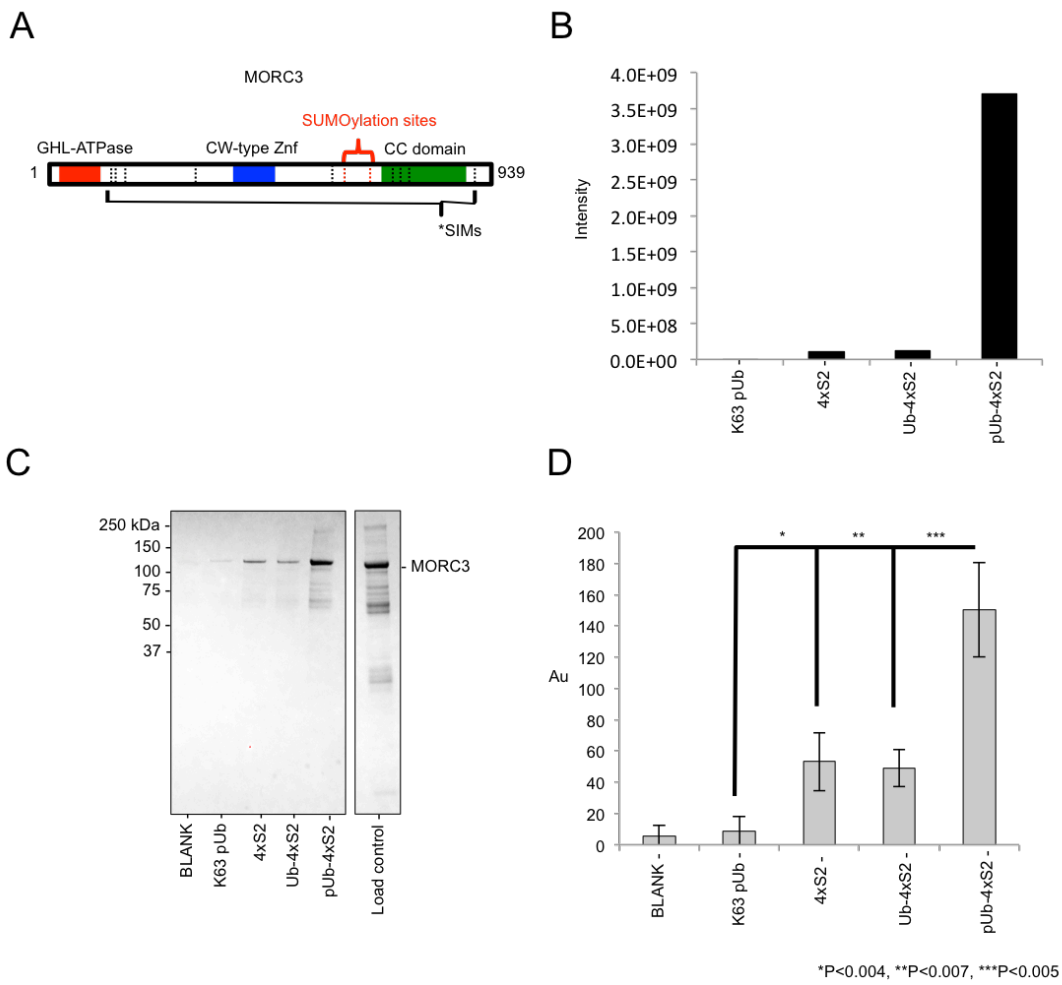


Figure 3.2-10 Morc3 a potential hybrid chain interactor.

A) Schematic diagram of MORC3 full length protein. N-terminally located GHL-ATPase domain represented as red box. CW-type Zinc finger represented as blue box. C-terminally located coiled coil domain (CC) represented as green box. Postion of potential SIMs represented by dashed black lines. Known SUMOylation sites at K597 and K650 represented by dashed red lines.

B) Combined intensities reproted by MaxQuant for Morc3 minus intenistiies reported for blank resin from hybrid chain pull down experiment (**Figure 3.2-7**)

C) SDS-PAGE analysis of hybrid chain pull down assay with recombinant full length Morc3. 50µl of each resin type used in pull down was incubated separately with 10µM of bacterially produced recombinant full length Morc3 overnight, before being eulated from the resins with SDS loading buffer. Load control represent 20% of total protein incubated with resin. SDS-PAGE gel is representative of 3 independent experiments with full length Morc3 used in pull down stated above and resin type used in pull down below.

D) Quantification of SDS-PAGE analysis of hybrid chain pull down assay with recombinant Morc3. Pull down experiment was carried out in triplicate. ImageJ was used to quantify Morc3 comassie stained bands for each resin. Error bars represent SEM. P values represent the

difference in average intenisties of 3 repeated experiments between * K63 pUb and pUb-4xSUMO2, ** 4xSUMO2 and pUb-4xSUMO2, and *** ub-4xSUMO2 and pUb-4xSUMO2.

To try and elucidate any domains of MORC3 that may be required for its apparent affinity for SUMO-ubiquitin hybrid chains, four truncated fragments of MORC3 were generated (**Figure 3.2-11, A**). These fragments were structured around the potential SIMs and known domains of MORC3. Fragment A contained the N-terminus of MORC3 that included four of the potential SIM domains plus the GHL-ATPase domain. Fragment B contained the GHL-ATPase domain, four SIMs but also included the CW-type Zinc finger domain. Fragment C contained the C-terminal of MORC3 including the known SUMOylation sites, the coiled coil domain, the CW-type zinc finger and five potential SIMs. Fragment D contained the C-terminal, coiled coil domain, SUMOylation sites and five SIM domains. These fragments of MORC3 were then tested in the SUMO-ubiquitin hybrid chain pull-down assay. Fragments A and B were found to be most abundant in the elutions from the 4xSUMO2 and Ub-4xSUMO2 resin types. Both fragments were found to be less abundant in the elutions from the pUb-4xSUMO2 and K63 pUb resin types. Fragments C and D were found to be present at similar levels in the 4xSUMO2, Ub-4xSUMO2, and pUB-4xSUMO2 and at lower levels in the K63 pUb resin types. The presence of potential SIM domains in all of the fragments would allow for the affinity towards SUMO seen in these pull down assays however none of the fragments show the great increase in affinity for the pUb-4xSUMO2 resin type seen in the full length protein. This could suggest that the affinity for SUMO-ubiquitin hybrid chains seen in the full length protein is not based on a discreet sequence domain of MORC3 but is perhaps related to the 3D structure of the full length protein. MORC3 clearly shows affinity for SUMO-ubiquitin hybrid chains however these MORC3 fragments have been unable to serve as tools to identify a discreet domain that defines such an affinity.

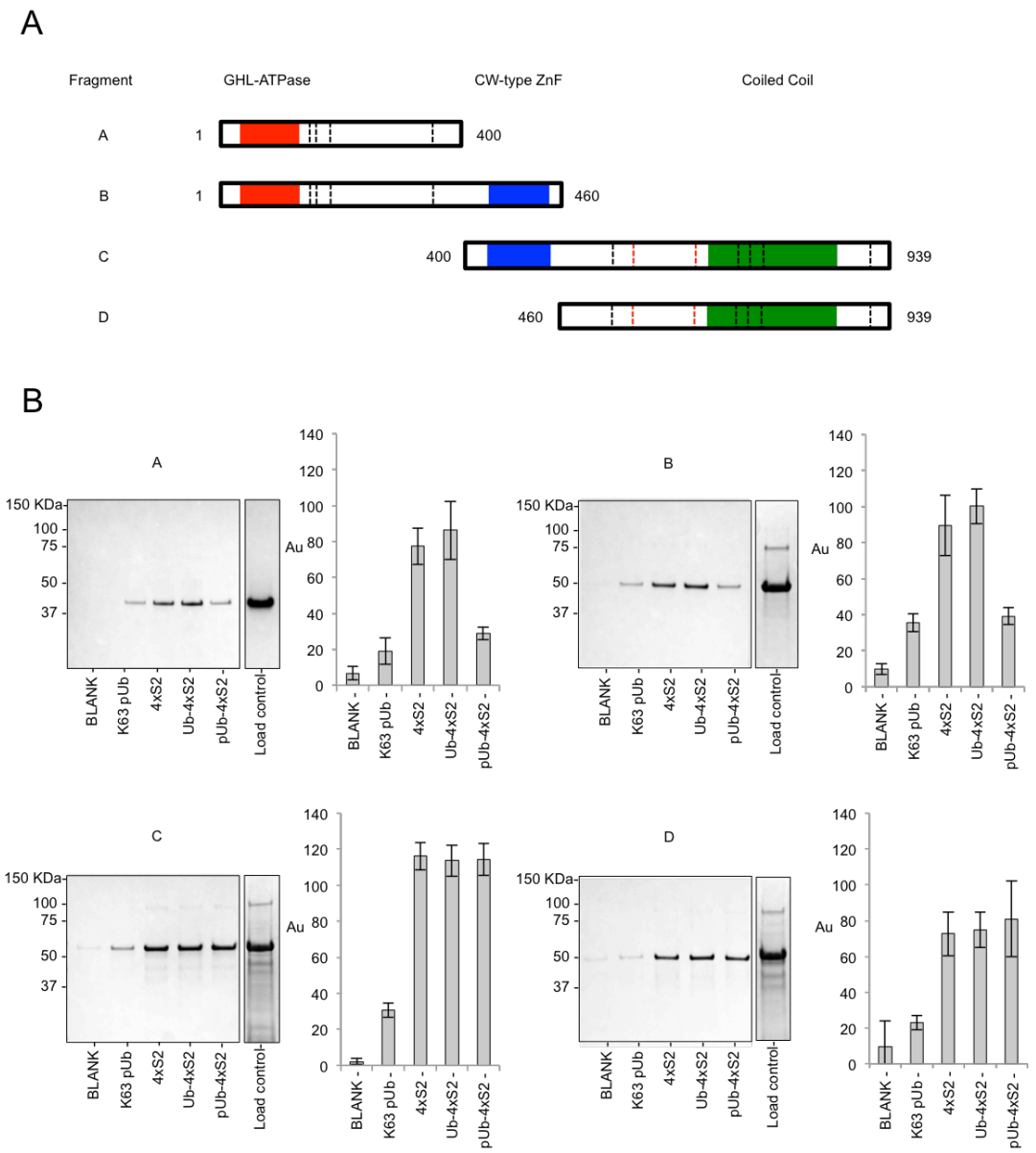


Figure 3.2-11 Dissecting the apparent interaction between MORC3 and hybrid chains of SUMO and ubiquitin.

A) Schematic diagram representing 4 truncated mutants of MORC3. Fragment A contain residues 1-400 encompassing the GHL-ATPase domain (red) and 4 potential SIMs (dashed black lines). Fragment B contains residues 1-460 encompassing the GHL-ATPase domain, 4 potential SIMs and the CW-type zinc finger domain (blue). Fragment C contains residues 400-939 encompassing the CW-type zinc finger 5 potential SIMs and coiled coil domain (green). Fragment D contains residues 460-939 encompassing the coiled coil domain and 5 potential SIMs.

B) Quantification of hybrid chain pull down assay with Morc3 fragments. Pull down assays were carried out with 50µl of each resin type and 10 µM of stated MORC3 fragment. Load controls represent 20% of total protein incubated with resins. SDS-PAGE gels are representative of 3 independent experiments with MORC3 fragment used in pull down stated

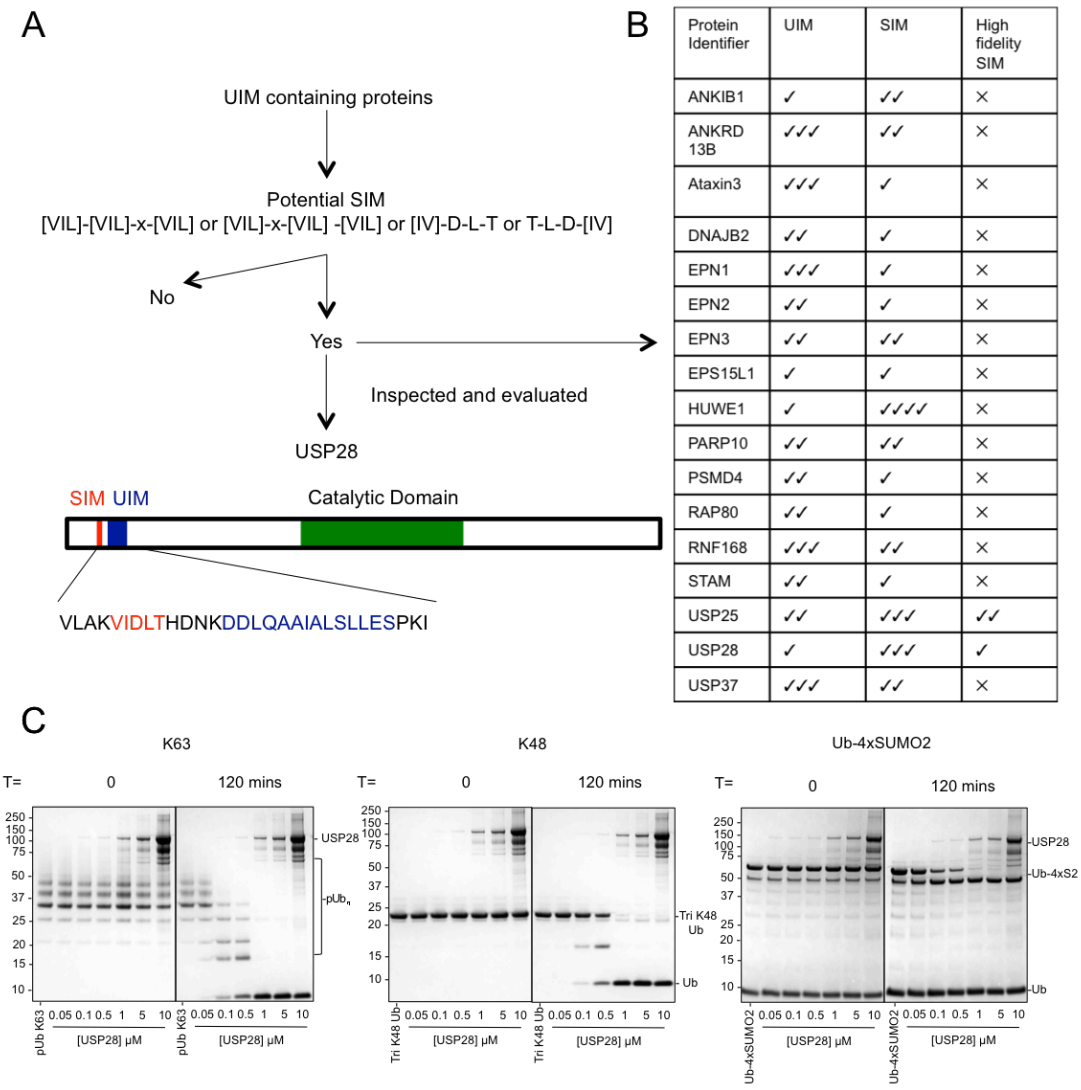
above and resin type used in pull down below. ImageJ was used to quantify MORC3 fragment coomassie stained bands for each resin. Error bars represent SEM.

3.2.5 A bioinformatic approach to identifying proteins with the potential to interact with SUMO-ubiquitin hybrid chains

As the SUMO interacting motifs (SIM) and a variety of ubiquitin binding domains, including the ubiquitin interacting motif (UIM) are small definable sequence specific motifs, it was postulated that by identifying proteins carrying both SIMs and UIMs, it could be possible to identify proteins with the ability to interact with SUMO-ubiquitin hybrid chains.

UIMs are thought to be general ubiquitin binding motifs and due to the presence of a duplet of UIMs found in RAP80 (Hofamnn & Falquet 2001, Sobhian et al., 2007), all human proteins known to contain UIMs were searched for SIMs. A list containing all human proteins thought to contain UIMs corresponding to "L-x-x-A-x-x-L-S-x-x-Ac", where "Ac" is an acidic residue, was obtained from Hofamnn & Falquet 2001 deposited at smart.embl.de. This list contains around 90 human proteins. The sequence data for each of these proteins was gained from UNiProt before each sequence was searched for the presence of potential SIM domains. The SIM domain searches were carried out using the PATTINPROT web tool at http://perso.ibcp.fr/gilbert.deleage/Cours/BIOINFO_NPS@_pattinprot.html. Canonical SIMs were searched for using the pattern values "[VILFY]-[VILFY]-x-[VILFY]" or "[VILFY]-x-[VILFY]-[VILFY]", were x corresponds to any amino acid. Search similarity criteria was set to; no mismatch=100%, this allowed no sequence to be reported that deviated form the queried sequence. All reported SIMs were then checked by eye for other sequence assets associated with SIM domains such as patches of negatively charged amino acids immediately up or down stream of the SIM domains or the presence of phosphorylated serine residues near the motif. Any SIM domains that met these criteria were then recorded. PATTINPROT was also used to query the UIM containing protein list for high fidelity SIM type. High fidelity SIMs were

searched for using the pattern value [VILFY]-[VI]-D-L-T or T-L-D-[VI]-[VILFY], search similarity criteria was again set to; no mismatch=100%. From the list of UIM containing proteins only two proteins gave positive hits when searched for the high fidelity SIM type, USP25 and USP28 (**Figure 3.1.12 B**). USP25, a ubiquitin protease, belonging to the USP subfamily has previously been reported to contain two high fidelity SIM domains and a doublet array of UIMs (Meulmeester et al., 2008, Demuc et al 2009). However these SIM domains are reported to be involved in the recruitment of a Ubc9-SUMO thioester and subsequent SUMOylation of a lysine situated within the UIM domains of USP25, impairing its non-covalent interaction to poly-ubiquitin and decreasing its deubiquitinase activity (Meulmeeester et al 2008). Therefore, it is unlikely that the SIM and UIM domains of USP25 would be involved in interacting with hybrid chains of SUMO and ubiquitin. USP28, a protease closely related to USP25, possesses one high fidelity SIM and one UIM in a tandem domain within its N-terminal region (**Figure 3.1.12 A, lower)**. USP28 has previously been implicated in the DNA damage response and is known to be an interacting partner of 53BP1, making USP28 an appealing target for further study (Zang et al., 2006, Popov et al., 2007). As such, full length USP28 was expressed and purified. The deubiquitinase activity of USP28 was then tested against recombinant K63-linked pUb, K48-linked tri Ub, and Ub-4xSUMO2. USP28 efficiently reduced K63-linked pUb and K48-linked tri Ub to mono-ubiquitin after 120 minutes at concentrations of 1 μM and above. Interestingly in the context of SUMO-ubiquitin hybrid chains, USP28 was able to remove the ubiquitin moiety from Ub-4xSUMO2 at concentrations of 1 μM and above (**Figure 3.1.12, C right hand panel**). USP28 therefore has the potential to both interact with SUMO ubiquitin hybrid chains and act to remove the ubiquitin moiety from the hybrid chain via its deubiquitinase activity.



Protein Identifier	UIM	SIM	High fidelity SIM
ANKIB1	✓	✓✓	×
ANKRD 13B	✓✓✓	✓✓	×
Ataxin3	✓✓✓	✓	×
DNAJB2	✓✓	✓	×
EPN1	✓✓✓	✓	×
EPN2	✓✓	✓	×
EPN3	✓✓	✓✓	×
EPS15L1	✓	✓	×
HUWE1	✓	✓✓✓✓	×
PARP10	✓✓	✓✓	×
PSMD4	✓✓	✓	×
RAP80	✓✓	✓	×
RNF168	✓✓✓	✓✓	×
STAM	✓✓	✓	×
USP25	✓✓	✓✓✓	✓✓
USP28	✓	✓✓✓	✓
USP37	✓✓✓	✓✓	×

Figure 3.2-12 Identifying potential hybrid chain interacting proteins using primary sequence data.

A) Experimental flow diagram. List of UIM containing proteins from smart.embl.de subsequently search for SIM regions via primary sequence search tool PATTINprot. Candidates were then inspected by eye before being taken forward for experimental work.

B) List comprises UIM containing human proteins found to also contain a minimum of one viable SIM domain. Number of ticks represents number of either know UIMs, or potential and viable SIMs, or known SIM domains

C) USP28 deubiquination assay. Concentration of USP28 [0-5 µM] shows deubiquitase activity against 10 µM; K63 linked poly ubiquitin (K63), K48 linked poly ubiquitin (K48), and mono ubiquitylated poly SUMO (Ub-4xSUMO2)

The potential high fidelity SIM and UIM domains located within the N-terminal of USP28 were then investigated to determine whether they could facilitate an interaction between USP28 and SUMO-ubiquitin hybrid chains. Three N-

terminal variants of USP28 were generated; SIM UIM which contained the putative high fidelity SIM and UIM domains, SIM ΔUIM containing only the high fidelity SIM, and MutSIM UIM that contained a mutated SIM and wild type UIM (**Figure 3.2-13 A**). These N-terminal fragments of USP28 were then tested in pull-down assays with the K63 pUb, 4xSUMO2, Ub-4xSUMO2, and pUb-4xSUMO2 resins to identify the affinity of each fragment to poly SUMO, poly-ubiquitin and SUMO-ubiquitin hybrid chains. The SIM UIM fragment shows affinity for both the 4xSUMO2 and K63 pUb, but shows higher affinity for the hybrid chain type resins Ub-4xSUMO2 and pUb-4xSUMO2. The SIM ΔUIM fragment shows affinity for the SUMO containing resin types 4xSUMO2, Ub-4xSUMO2, and pUb-4XSUMO2. A small amount of SIM ΔUIM was present in the eluate from the K63 pUb resin. The ΔSIM UIM fragment showed affinity for the ubiquitin containing resin types, K63 pUb, Ub-4xSUMO2, and pUb-4xSUMO2. Thus the SIM UIM tandem domain of USP28 appeared to have the potential to interact non-covalently with SUMO via the high fidelity SIM domain and with ubiquitin via the UIM domain and thus with SUMO-ubiquitin hybrid chains.

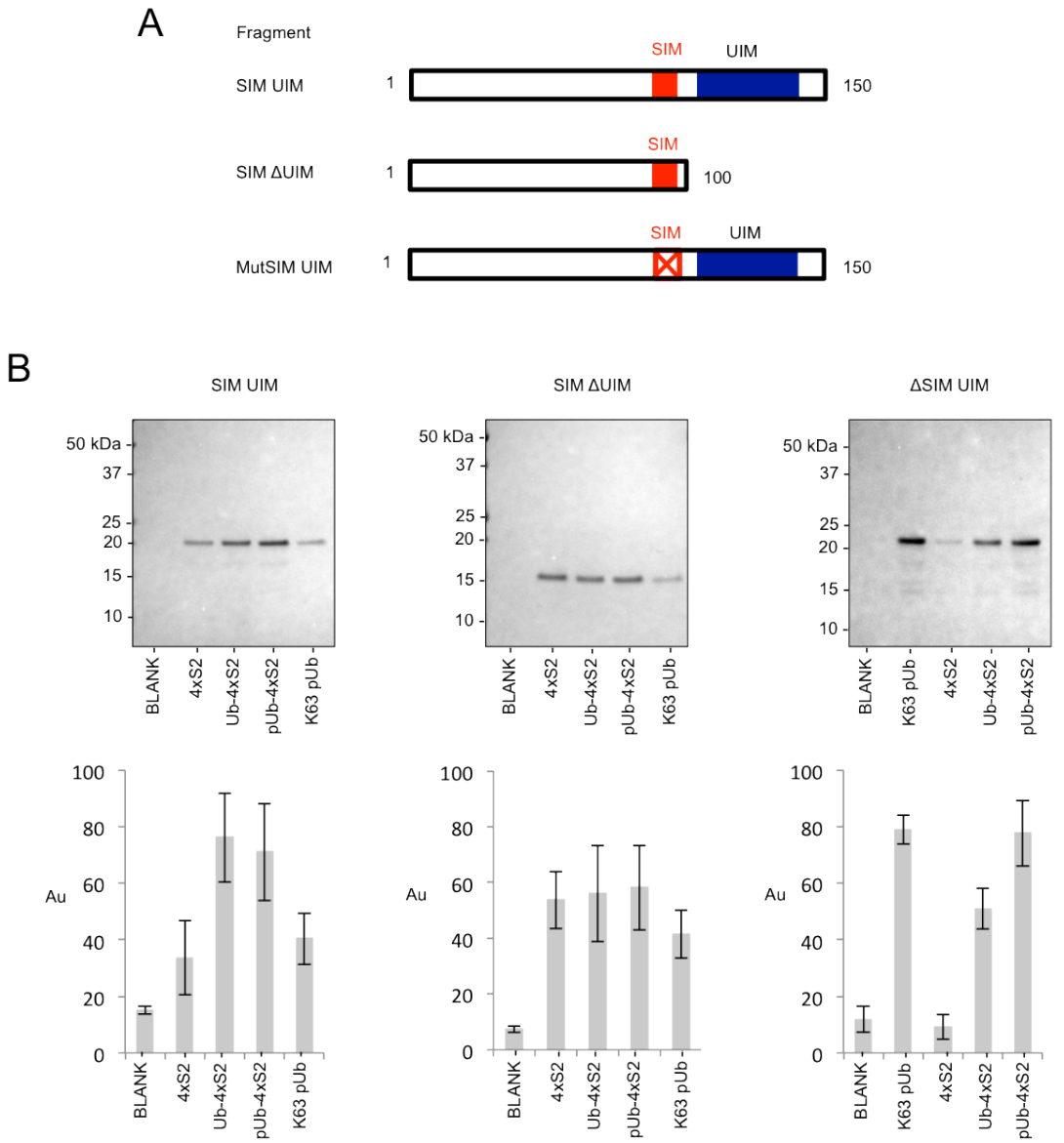


Figure 3.2-13 N-terminally located SIM UIM domains of USP28 show affinity for SUMO and ubiquitin.

A) Schematic diagram of USP28 showing the wild type SIM UIM N-terminal residues 1-150 including wt SIM and UIM domains (top); SIM ΔUIM fragment residues 1-100 (middle); and MutSIM UIM residues 1-150 with SIM domain mutated from VIDLT to AAAAT (bottom).

B) Quantification of USP28 N-terminal fragment in pull down assay SDS-PAGE analysis.

However during further investigation of USP28, several papers were published suggesting that in a similar fashion to the SIM domains of USP25, the SIM domain of USP28 acted in the recruitment of the SUMO-Ubc9 thioester resulting in the SUMOylation of a lysine residue near the UIM of USP28, impairing the non-covalent interaction with ubiquitin (Knobel et al.,

2014, Zhen et al., 2014). A further UBA domain in the N-terminal of USP28 was also confirmed, and although USP28 is recruited to sites of DNA damage it seemingly plays only a minor role in any response to the damage (Wu et al., 2013, Knobel et al., 2014). As such, study of USP28 and any interaction with SUMO-ubiquitin hybrid chains was not taken any further.

3.3 Discussion

Post-translational modification by the Ubl family members SUMO and ubiquitin are critically important for many cellular processes. Once thought to be distinct pathways, it has now become clear through the action of STUbLs such as RNF4 that a direct interaction exists between the SUMOylation and ubiquitylation systems. This is most clearly apparent during the arsenic-induced degradation of PML where by the critical action of RNF4 results in the decoration of target proteins with SUMO-ubiquitin hybrid chains (Tatham et al., 2008). Identifying proteins that interact with these hybrid chains will help to elucidate the biological role these chains play in the cell. Recently published literature suggests that proteins such as RAP80 can bind with high affinity to SUMO-ubiquitin hybrid chains. Here for the first time an affinity chromatography coupled to mass spectrometry based approach is utilized to elucidate novel protein interactions with SUMO-ubiquitin hybrid chains.

3.3.1 Generating SUMO-ubiquitin hybrid chains

Firstly, hybrid chains of SUMO and ubiquitin had to be created. Two distinct types of hybrid chains were designed with the known action of the STUbL RNF4 taken into consideration. A poly-SUMO chain with a single N-terminal ubiquitin modification and a second hybrid chain type consisting of poly SUMO N-terminally modified with ubiquitin that was subsequently poly-ubiquitylated with K63 linked ubiquitin chains. As the STUbL RNF4 contains 4 SIMs, both hybrid chain types contained a poly-SUMO moiety comprised of four SUMO2 proteins. These four SUMO2 moieties have been shown to recruit and activate RNF4 *in vivo* (Tatham et al., 2008, Rojas-Fernandez et al., 2014). In this case, the four SUMO2 moieties were expressed as a single linear head-to-tail fusion protein. This allowed the poly-SUMO portion of each of the three poly-SUMO containing resins to be standardized and produced quickly, while still retaining the ability to interact with RNF4 (**Figure 3.2-2**). Hybrid chains of poly-SUMO and poly-ubiquitin could theoretically be made from any number of possible SUMO and ubiquitin linkages, allowing a vast number of variants of hybrid chain types to exist (see **3.1.2**). Recent evidence

suggests that the ubiquitin E2 conjugating enzyme Ube2W can act to N-terminally mono-ubiquitylate SUMO2 (Tatham et al., 2013). This can act as a primer for poly-ubiquitylation of SUMO chains by RNF4 in conjunction with other ubiquitin E2 conjugating enzymes. Currently no receptors are known for this N-terminally modified type of SUMO-ubiquitin hybrid chain. K63 linked poly-ubiquitin was chosen as the poly-ubiquitin element of the poly-SUMO poly-ubiquitin hybrid chain due to the known role of K63 linked poly-ubiquitin in signalling DNA repair pathway and as a recent evidence from our lab suggests depletion of RNF4 in cells results in a significant decrease in the K63 linked poly-ubiquitin chains present at the sites of DNA damage (Yin et al., 2012). Having both hybrid chain types allowed a differentiation to be made between proteins that preferentially interact with N-terminally ubiquitylated poly-SUMO such as TULP4 or proteins that require poly-ubiquitylated poly-SUMO such as RAP80 (**Figure 3.2-7, C**).

Although the number of theoretical SUMO-ubiquitin hybrid chain types is large, the hybrid chain types used in this experiment are relevant in the context of the known action of the STUbL RNF4 and proved capable of elucidating a number of novel targets for SUMO-ubiquitin hybrid chain interacting proteins.

3.3.2 USP2 CaD/SENP1 specific elution

A specific elution step was also investigated to help reduce elution complexity and improve protein specificity identification. The elution based around sequentially adding ubiquitin and SUMO proteases, USP2 CAD and SENP1 respectively, allows the hybrid chains to be disassembled in manner that allows proteins interacting with the ubiquitin element of the hybrid chains to be separated from the SUMO interacting proteins of the hybrid chains. A final denaturing elution with SDS/LDS sample buffer acts as the final step of the elution process removing any proteins still bound to the partially or undigested hybrid chains. This elution protocol is effective in giving information about the potential binding affinities of some proteins bound to the

hybrid chains. For example in the case of C5orf25, the USP2 CaD elution suggested that the ubiquitin moieties of the SUMO-ubiquitin hybrid chain containing Ub-4xSUMO2 and pUb-4xSUMO2 resins were contributing to the affinity C5orf25 showed toward those resin types. C5orf25 was released from those resins after USP2 CaD elution but not from the SUMO only containing 4xSUMO2 resin (**Figure 3.2-9**). This USP2 Cad elution data along with the far greater intensity of C5orf25 found in the eluate from the pUb-4xSUMO2 resin suggested that C5orf25 was indeed interacting with both the SUMO and ubiquitin elements of the hybrid chains. However, additional information could not always be gathered via the specific elution as can be seen for RNF168, where RNF168 is only detected in significant amounts after the denaturing elution step from the pUb-4xSUMO2 resin (**Figure 3.2-9**). This is could be due to RNF168 and other proteins binding tightly to the hybrid chains blocking USP2 CAD and SENP1 from degrading the hybrid chains.

3.3.3 An assessment of the approach to detect SUMO-ubiquitin hybrid chain interacting proteins via affinity baits coupled with high resolution mass spectrometry.

Of the 2197 proteins identified by mass spectrometry after the SUMO-ubiquitin hybrid chain pull down assay, 30 were characterised as showing specificity for SUMO-ubiquitin hybrid chains with the remaining 122 proteins observed to show a specific affinity for K63 linked poly-ubiquitin or poly-SUMO. RAP80 was identified as expected as a SUMO-ubiquitin hybrid chain interacting protein along with five other members of the BRCA 1A complex (**Figure 3.2-8**) (Guzzo et al., 2012). This suggests that these proteins may have been pulled down as a complex. Other known members of the BRCA 1A complex such as Abraxas, BRCC36 and BRE are predicted to also have SIMs domains as well as UBDs, so it is very possible that these proteins act together with RAP80 to establish an interaction between the BRCA 1A complex and SUMO-ubiquitin hybrid chains (**Figure 3.2-8**). It may also be possible that members of the BRCA 1A complex are themselves modified by hybrid chains of SUMO and ubiquitin which then in turn aids the formation and

association of the BRCA 1A complex at the sites of DNA damage. Of the remaining interesting candidate proteins three were taken for further validation RNF168, C5orf25 and MORC3.

The ubiquitin E3 ligase RNF168 has previously been shown to have a role in the DNA damage response recruiting repair proteins to the sites of double strand breaks (Doli et al., 2009; Pinato et al., 2009; Mattiroli et al., 2012). It is thought the RNF168 is recruited to DNA damage by binding to K63 linked poly-ubiquitylated histones via three UBDs- 1UIM and 2 MIUs (Pinato et al., 2009). The action of RNF168 and another RING type ubiquitin E3 ligase, RNF8, is then thought to act to recruit other DDR proteins to the site of damage including RAP80 (Mattiroli et al., 2012). Sequence searches suggest four potential SIMs in RNF168, with two of these SIMs being located proximal to the second UIM of RNF168. RNF168 interaction with poly-SUMO poly-ubiquitin hybrid chains was backed up with data from recombinant RNF168 tested in the SUMO-ubiquitin hybrid chain pull down assay (**Figure 3.2-9, H**). In accordance with the data analysed from the mass spectrometry experiment, the recombinant RNF168 protein was also found predominantly in the elution from the pUb-4xSUMO2 resin type. In agreement with the current literature recombinant RNF168 also appeared in the eluate from the poly-ubiquitin only, K63 pUb resin type (Doli et al., 2009). Taking the literature and the data presented here into account, RNF168 appears to have the potential to interact with SUMO-ubiquitin hybrid chains. However, further validation of the specific domains related to this apparent affinity for hybrid chains would be required. RNF4 is known to be present at the sites of DNA damage, so it is possible that RNF168 could be recruited to these sites through the formation of RNF4-dependent SUMO-ubiquitin hybrid chains.

C5orf25 is suggested to have a role as a regulatory scaffold for the skeletal muscle-specific calpain, CAPN3 (Ono et al., 2013) but is not known to have any biological role involving SUMO or ubiquitin. C5orf25 contains two N-terminally located high fidelity type SIM domains (Hunter & Sun 2012). Unsurprisingly C5orf25 shows evidence of an interaction with poly-SUMO in the mass spectrometry data as it is found in the eluate from all of the poly-

SUMO containing resin types; 4xSUMO-2, Ub-4xSUMO-2 and pUb-4xSUMO-2. However, the greatest intensity of C5orf25 is found in the eluates from the pUb-4xSUMO2 resin. The USP2 CAD elution suggests that C5orf25 may interact with ubiquitin. This is also backed up by the data from testing the [^{35}S] Methonine labelled recombinant C5orf25 in the SUMO-ubiquitin hybrid chain pull-down assay. The bands present after elution from both hybrid chain types, Ub-4xSUMO-2 and pUb-4xSUMO-2 are more prominent than for 4xSUMO-2 alone suggesting that more C5orf25 bound the hybrid chains (**Figure 3.2-9**). Unlike RNF168, C5orf25 does not contain a defined ubiquitin binding domain. However, the presence of two high fidelity SIM domains and the data presented here suggest that C5orf25 may be a SUMO-ubiquitin hybrid chain interacting protein.

MORC3 was identified as a putative poly-SUMO poly-ubiquitin interacting protein in the mass spectrometry analysis. This apparent affinity for poly-SUMO poly-ubiquitin hybrid chains was also seen with recombinant MORC3 tested in the SUMO-ubiquitin hybrid chain pull-down assay. MORC3 belongs to a family of CW-type zinc finger containing proteins. Interestingly, MORC4, another member of this family was also identified as a potential hybrid chain binder by the mass spectrometry data. MORC3 and MORC4 share 34% sequence identity (Alignment using UniProt align, data not shown), showing large amounts of homology in the N-terminal regions of each protein. However, when MORC3 N-terminal fragments (**Figure 3.2-11, B, MORC3 A & B**) were tested in the SUMO-ubiquitin hybrid chain pull-down assay, the apparent affinity for the poly-SUMO poly-ubiquitin hybrid chain type shown with the full length protein was not apparent. This suggests that although MORC3 and MORC4 both appear to interact with hybrid chains of SUMO and ubiquitin, any discrete domain facilitating this apparent affinity is not located in the shared N-terminal regions of these proteins. The affinity for SUMO-ubiquitin hybrid chains does not appear to be generated by the shared CW-type zinc fingers of these proteins as removing the zinc finger from either the N-terminal fragment or the C-terminal fragment (**Figure 3.2-11, B, comparing Fragment A to B and C to D**) does not affect the affinity of MORC3 for any of the resin types used in the SUMO-ubiquitin hybrid chain pull-down assay.

This leads to the conclusion that the apparent affinity MORC3 shows for SUMO-ubiquitin hybrid chains does not appear to be facilitated by a discrete sequence domain of MORC3, but this affinity may be facilitated by folding of the full length protein.

3.3.4 An assessment of the approach to detect SUMO-ubiquitin hybrid chain interacting proteins via primary sequence searches

The strategy for identifying SUMO-ubiquitin hybrid chain interacting proteins via primary sequence searches involved taking a list of already identified UIM containing proteins and searching the sequences of these proteins for SIM domains using the online PATTINPROT search tool. To some extent this was successful, with 17 proteins previously identified to contain UIM domains also identified as containing SIM domains. Interestingly, RNF168, HUWE1, EPN1, EPN2, EPS15L1 and RAP80 were identified in both the list generated by the SUMO-ubiquitin hybrid chain pull down assay and by primary sequence searches. This adds confidence to the results gained from the mass spectrometry analysis of the SUMO-ubiquitin hybrid chain pull-down assay. USP28, a ubiquitin protease known to contain one UIM domain was shown also to contain one high fidelity type SIM domain. Interestingly the UIM and SIM domains were located in tandem, similar to the tandem SIM UIMUIM domain of RAP80. The N-terminal domain of USP28 containing the SIM and UIM showed the highest affinity for both SUMO-ubiquitin hybrid chain containing resin types, however, it also showed affinity for the poly-SUMO only resin type and the poly-ubiquitin only resin type. Primary sequence searches allowed proteins containing both ubiquitin binding domains and SIMs to be identified. However, care must be taken when exploring whether these proteins are indeed bona fide SUMO-ubiquitin hybrid chain interacting proteins. In the case of USP28, although proteins can contain both SUMO interacting and ubiquitin interacting sequence elements, this does not necessarily mean that the biological function of these sequence domains is to interact with SUMO-ubiquitin hybrid chains.

3.3.5 Conclusions

The strategies set out above have successfully identified a group of proteins with the potential to interact specifically with hybrid chains of SUMO and ubiquitin. A pull-down assay utilising hybrid chains of SUMO and ubiquitin was developed to allow for proteins to be identified that showed affinity specifically for SUMO, ubiquitin or SUMO-ubiquitin hybrid chains (**Figures 3.2-5 & 3.2-6**). The hybrid chain pull-down assay coupled with high-resolution mass spectrometry identified 30 proteins with specific affinity for SUMO-ubiquitin hybrid chains (**Figure 3.2-7**). A number of interesting candidates were taken for further investigation. Recombinant C5orf25 (**Figure 3.2-9, A-D**), RNF168 (**Figure 3.2-9, E-F**), and MORC3 (**Figure 3.2-10**) all showed similar affinity for SUMO-ubiquitin hybrid chains observed in the mass spec data. This gives credence to the hypothesis that SUMO-ubiquitin hybrid chains could form a new class of PTM; merging both the SUMO and ubiquitin modification systems. Further work is required however define the domains responsible for the apparent interactions of C5orf25 and MORC3 for SUMO-ubiquitin hybrid chains. The affinity of RNF168 for hybrid chains of SUMO and ubiquitin is hypothesised to be due to the presence of a tandem SIMSIM MIU domain located near the N-terminal region of the protein. Again further mutational studies will be required to show the importance of this SIMSM MIU domain in the context of RNF168 and SUMO-ubiquitin hybrid chains. The BRCA 1A complex component RAP80, a protein previously observed to have the potential to interact with hybrid chains of SUMO and ubiquitin (Guzzo et al., 2012) was identified in both the SUMO-ubiquitin hybrid chain pull-down assay (**Figures 3.2-7 & 3.2-8**) and the list of protein known to contain UIMs that also contain SIMs (**Figure 3.2-13, B**) making it the primary candidate for a bona fide SUMO-ubiquitin hybrid chains interacting protein. As such it shows great potential for further study to elucidate a potential biological role for SUMO-ubiquitin hybrid chains.

4 Exploring a role for SUMO-ubiquitin hybrid chains in the recruitment of RAP80 to the sites of DNA repair

4.1 Introduction

Identifying a number of proteins that showed affinity for hybrid chains of SUMO and ubiquitin is very interesting but without biological context it is hard to determine whether or not SUMO-ubiquitin hybrid chains and proteins capable of recognising said chains are relevant signalling/receptor mechanisms within the cell.

4.1.1 DNA damage response, RAP80 and the BRCA 1A complex

DNA lesions are highly toxic if not properly recognised and repaired. The presence of one unrepaired DNA double strand break (DSB) is sufficient to cause cell death or cancer (Jackson & Bartek 2009). Eukaryotic cells have developed a complex set of signalling cascades involving activation of cell cycle check points and recruitment of factors involved in the recognition, signalling, and repairing of DNA to the sites of damage, known collectively as the DNA damage response (DDR) (Jackson & Bartek 2009). DNA can be damaged in multiple ways and thus multiple systems are in place to repair these different types of DNA lesions, DSBs are the most cytotoxic of all DNA lesions. DNA DSBs initiate a signalling cascade mediated by one of the master DNA damage signalling protein kinases ataxia telangiectasia mutated (ATM) after recognition by the MRE11-RAD50-NSB1 (MRN) complex (Maser et al., 1997; Falck et al., 2005) (**Figure 4.1-1**). This results in repair of the DSBs by homologous recombination (HR) or non-homologous end joining (NHEJ) (Lundin et al., 2002). HR is an error-free process utilising homologous sequence from sister chromatids to repair breaks. Due to the requirement for sister chromatid sequence homology, HR can only function during the S and G2 phases of the cell cycle- after replication has occurred. NHEJ is an error-prone process but does not require sister chromatid sequence homology and thus is not restricted to specific cell cycle phases (Lundin et al., 2002).

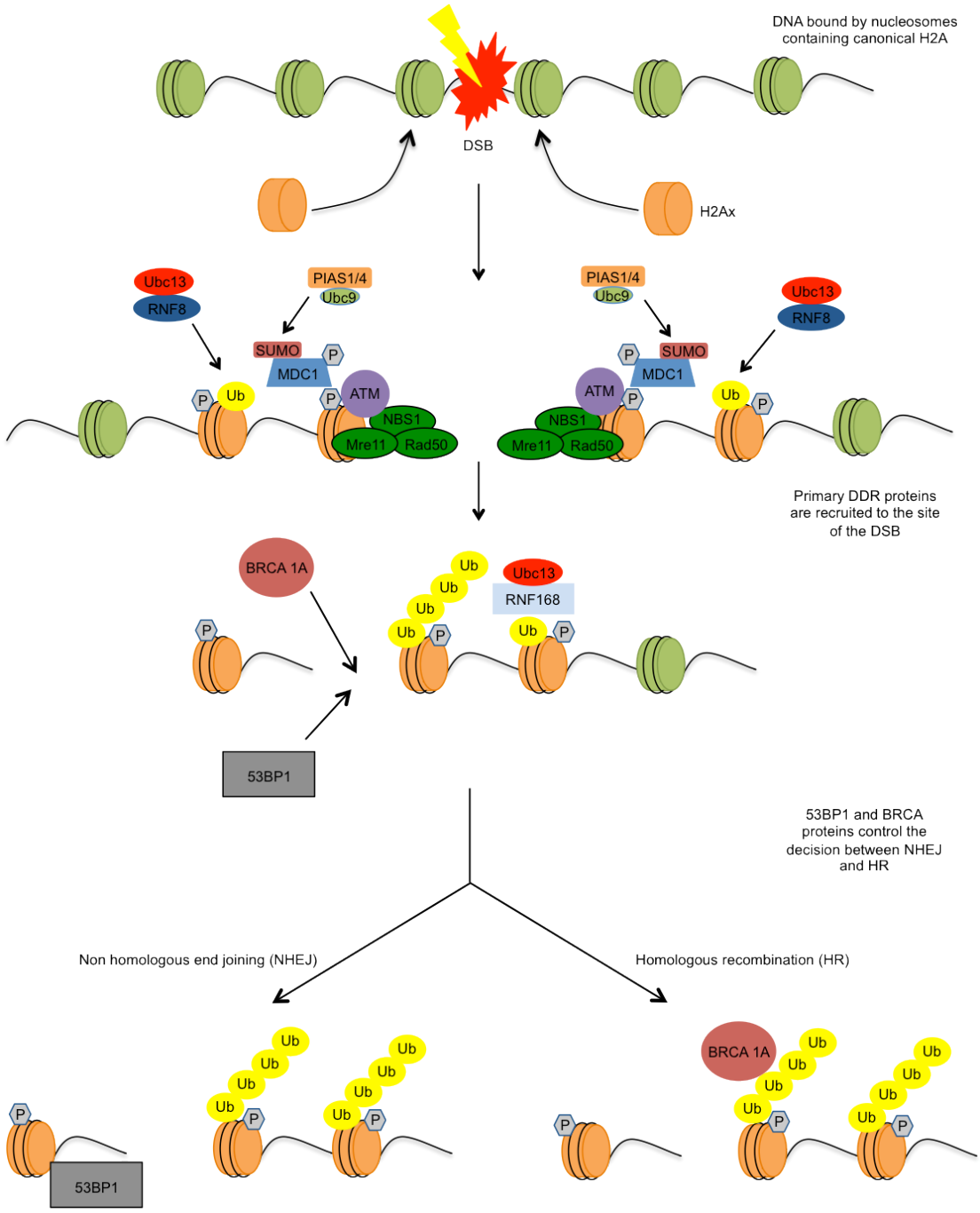


Figure 4.1-1 The DNA damage response to DSBs.

After a DSB, Histone H2Ax replaces the canonical H2A at the site of damage. The MRN complex is recruited which in turn recruits ATM. ATM phosphorylates (P) and recruits MDC1 to the sites of damage. MDC1 is also SUMOylated at the sites of damage. The phosphorylation of MDC1 then engages RNF8. RNF8 is then thought to catalyse the mono-ubiquitylation of targets near the site of damage. This mono-ubiquitylation then recruits RNF168 resulting in the poly-ubiquitylation of targets at the sites of damage. BRCA and 53BP1 are then recruited to the DSB, initiating either the HR or NHEJ DNA repair pathways.

The decision between HR and NHEJ is mediated essentially by the recruitment of two important factors BRCA1 and 53BP1 (**Figure 4.1-1**) (Noon & Goodarzi 2011; Li & Greenberg 2012). BRCA1 and 53BP1 are thought to perform opposing functions with respect to DNA end resection, which is vital for HR. 53BP1 acts to restrict DNA end resection and thus promotes repair by NHEJ (Noon & Goodarzi 2011), whereas BRCA1, through mechanisms that are not full understood initiates DNA end resection (Li & Greenberg 2012). After DNA end resection has been initiated NHEJ is suppressed and DSBs are repaired via HR (Chapman et al., 2012).

BRCA1 is a known component of multiple protein complexes (Yu et al., 1998; Cantor et al., 2001; Wang et al., 2007). During the activation of HR the BRCA 1A complex is thought to be located at the sites of damage (Wang et al., 2007). The BRCA 1A complex is composed of at least Abraxas, BRCC36, BRE, Merit40 and RAP80 (Wang et al., 2007). Initially a key finding was that RAP80 mediated the association of the BRCA 1A complex with sites of DNA damage, by recognising and binding to K63 linked poly-ubiquitin chains that were thought to be anchored on K119 of phosphorylated histone H2Ax (phosphorylated at S139, commonly referred to as yH2Ax (Burma et al., 2001)) which marks sites of DNA damage (Sobhian et al., 2007;Wu et al., 2007). This is because the double UIM containing N-terminal region of RAP80 was initially observed to be required for the correct association and retention of BRCA1 at the sites of DSBs (Sobhian et al., 2007). This key finding and others that followed helped to cement the role of ubiquitylation in DDR.

4.1.2 SUMO, ubiquitin and the DNA damage response

Around the time that K63 linked poly-ubiquitin was identified as an important factor in the recruitment of BRCA1 to the sites of DSBs, two RING containing ubiquitin E3 ligases were identified as key factors in the DDR, RNF8 and RNF168 (Wang & Elledge 2007). RNF8 and RNF168 work in tandem with the ubiquitin E2 conjugating enzyme Ubc13 and together act to catalyse the addition of K63 linked poly-ubiquitin chains to histones and most

probably other proteins involved in the DDR (Stewart et al., 2009; Doil et al., 2009). RNF8 recruitment to the sites of DSBs is facilitated by an interaction with phosphorylated MDC1. MDC1 is recruited to site of DNA damage by the presence of yH2Ax and is then phosphorylated by ATM (**Figure 4.1-1**) (Stucki et al., 2005). RNF8 acts to catalyse the mono-ubiquitylation of substrate proteins. This mono-ubiquitin mark is then recognised and modified on K63 by RNF168 resulting in the generation of K63 linked poly-ubiquitin, interestingly RNF168 may act as a negative regulator in the RNF8 response and thereby acting to limit excessive K63 linked poly-ubiquitin signalling (Gudjonsson et al., 2012). Initially, It was this K63 linked poly-ubiquitin signal that was thought to be the primary recruitment mechanism for other DDR proteins including 53BP1, RAP80, and thus the BRCA 1A complex to sites of DNA damage (Lukas et al., 2011). However more recent evidence has suggested a role of the STUbL RNF4 at the sites of DNA damage and importantly to the recruitment of the BRCA 1A complex to the sites of DNA damage via a RAP80 mediated interaction with SUMO-ubiquitin hybrid chains, suggesting SUMO plays a key role in the DDR (**Figure 4.1-2**) (Guzzo et al., 2012).

This was not the first observation of SUMO in the DDR however. SUMO was identified to have a role in the DDR in with the observation that in the yeast *S. cerevisiae* Rad52 is SUMOylated (Torres-Rosell et al., 2007). Rad52 is an important factor in DNA repair, promoting the annealing of complementary single stranded DNA. The SUMOylation of Rad52 was dependent on Ubc9 and the E3 Siz2, a protein highly related to the human PIAS protein family. SUMOylation ablates the activity of Rad52 by functioning to exclude it from the nucleus (Torres-Rosell et al., 2007). All mammalian SUMO isoforms have been observed at DSBs along with Ubc9 and the PIAS SUMO E3 ligases (Galanty et al., 2009; Morris et al., 2009). SUMO, PIAS1 and PIAS4 clearly have critical roles in the signalling of repair proteins at these sites. Knockdown of either PIAS1 or PIAS4 impairs both HR and NHEJ by decreasing the levels of RNF168, 53BP1 and BRCA1 at these lesions, thus impairing DNA repair (Morris et al., 2009). Interestingly, knockdown of RNF8 and RNF168 does not inhibit the accumulation of PIAS1 and PIAS4 to

sites of damage. SUMO levels were reduced but not ablated, suggesting that DDR factors recruited to DSBs both prior and post RNF8/RNF168 activity can be SUMOylated.

4.1.3 SUMO-ubiquitin hybrid chains at the sites of DNA damage

STUbLs were first connected to the DDR with the still not fully understood observation that HR in *S. cerevisiae* is regulated by the heterdimeric Slx5-Slx8 (Yang et al., 2006; Uzunova et al., 2007). A hand full of complementary papers were more recently published on the role of the human STUbL RNF4 in HR suggesting its recruitment to DSBs was facilitated by the SUMOylation of key DDR proteins including MDC1 (Galanty et al., 2012; Luo et al., 2012; Yin et al., 2012). The partial knock down of RNF4 resulted in a decrease in the presence of K63 linked poly-ubiquitin at DSBs (Yin et al., 2012). These observations are in line with the model suggested by Guzzo and colleagues for RNF4-dependent SUMO-ubiquitin hybrid chain recruitment of RAP80 to the sites of DSBs (Guzzo et al., 2012), whereby, RNF4-generated K63 linked poly-ubiquitylated SUMO is required for the recruitment of the BRCA 1A complex, mediated by the tandem SIM UIMUIM domain of RAP80, to the sites of DNA damage (**Figure 4.1-2**). If knocking down RNF4 lowers the levels of K63 linked poly-ubiquitin (Yin et al., 2012), and thus then SUMO2 modified by K63 linked poly-ubiquitin at the sites of DNA damage this could explain why the BRCA 1A complex and thus BRCA1 is not recruited to the sites of DNA damage under these conditions (Guzzo et al., 2012). Although ubiquitylated SUMO has been observed in cells (Tammsalu et al., 2014) and clearly a level of cooperativity exists between the SUMO and ubiquitin conjugation pathways in the DDR, with multiple substrates for each pathway known, its is unclear whether this cooperativity extends to any type of SUMO-ubiquitin hybrid chains forming at the sites of DSBs. No potential substrates have yet been identified to anchor SUMO-ubiquitin hybrid chains at the sites of DSBs. However, yH2Ax is the primary candidate for RAP80 recruitment to sites of damage via the poly-ubiquitylation of K119. It is interesting to note that recently SUMO2 has also been observed to modify K119 of histone H2Ax leaving open the possibility that Histone H2Ax

could anchor a SUMO-ubiquitin hybrid chain, but it is important to note that this observation was made not under specific genotoxic stress but after heat shock (Tammsalu et al., 2014).

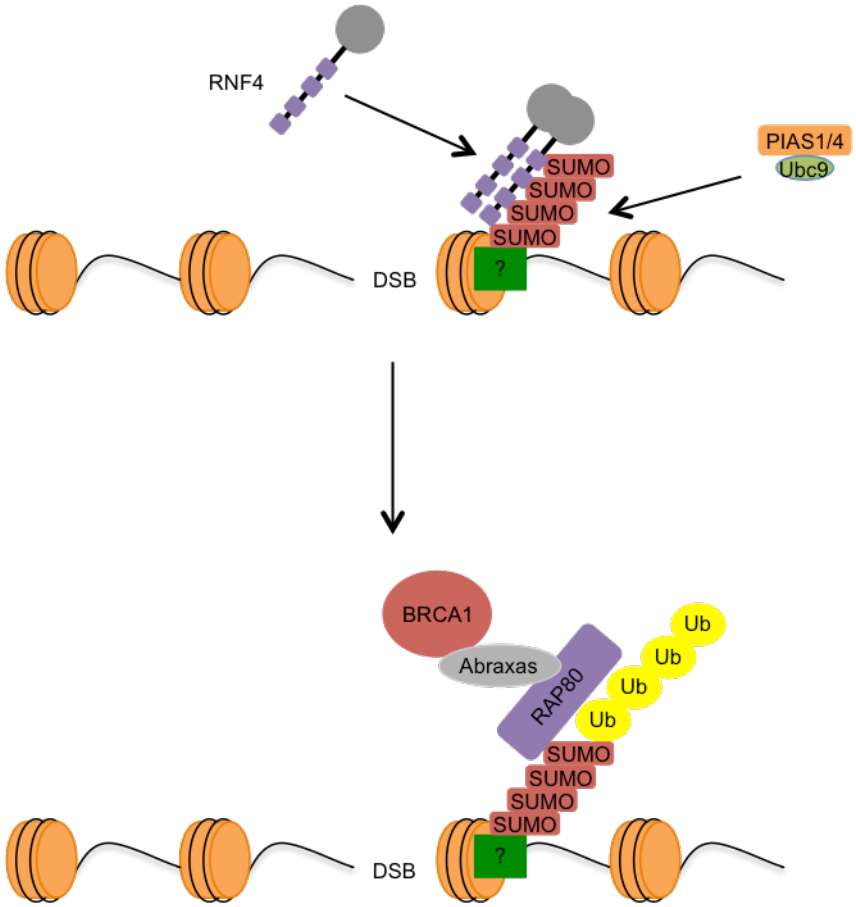


Figure 4.1-2 RNF4-dependent RAP80 mediated recruitment of BRCA1.

After the DDR to a DSB has been initiated, an as yet unidentified target protein at the site of the DSB is SUMOylated. This SUMOylated target substrate is then targeted by the STUbL RNF4 resulting in the K63 linked poly-ubiquitylation of the previously SUMOylated substrate. This SUMO-ubiquitin hybrid chain is then recognised by the BRCA 1A complex component RAP80. RAP80 then mediates the recruitment of BRCA1 (Guzzo et al., 2012).

Clearly the roles of SUMO and ubiquitin in the DDR are very complex and not yet fully understood. SUMO-ubiquitin hybrid chains acting as a signalling mechanism at the sites of DSBs is a compelling idea. Firstly, the specificity that a hybrid SUMO-ubiquitin signal would impose on the system could regulate a very specific response at these sites. Secondly, RNF4 and other essential enzymes from both the ubiquitylation and SUMOylation

pathways required to form such signals are known to be prevalent at these sites (Wang & Elledge 2007; Galanty et al., 2009; Morris et al., 2009; Galanty et al., 2012; Luo et al., 2012; Yin et al., 2012). As RAP80 has been observed previously to have affinity for hybrid chains (see **Figure 3.2-7),** and as RNF4 may have a role in the recruitment of the BRCA 1A complex to the sites of DSBs (Guzzo et al., 2012), the validity of RNF4-dependent SUMO-ubiquitin hybrid chain recruitment of RAP80 to the sites of DSBs during the DDR is worthy of further exploration.

4.2 Results

Having identified a number of novel potential SUMO-ubiquitin hybrid chain interacting proteins via firstly, the affinity chromatography coupled proteomic approach, and secondly, using a primary sequence searching based approach, a biological role for SUMO-ubiquitin hybrid chains was then investigated.

4.2.1 BRCA 1A complex component RAP80 shows an association for SUMO and ubiquitin in a cellular context

The mechanism by which the BRCA 1A complex is recruited to sites of DNA repair has been previously covered in the literature (Sobhian et al., 2007; Wang et al., 2007), however the exact mechanism by which RAP80 targets the BRCA 1A complex to sites of DNA repair is still under investigation. Here, microscopy and proteomic based approaches were used to investigate the potential role, if any, that hybrid chains of SUMO and ubiquitin could play in recruiting RAP80 and the BRCA 1A complex to the sites of DNA damage.

SUMO, ubiquitin and RAP80 are all known to be present at the sites of DNA repair during the DNA damage response (Sobhian et al., 2007; Wang et al., 2007; Doil et al., 2009; Morris et al., 2009). Determining that all three of these proteins are present at the same DDR foci at the same time is critical to understanding whether SUMO-ubiquitin hybrid chains could play a role in recruiting RAP80 to the sites of DNA repair. RAP80 has previously been shown to contain a tandem SIM UIMUIM domain in its N-terminal region (**Figure 3.1-2**) (Guzzo et al., 2012). RAP80 was identified as a hybrid chain interacting protein in the SUMO-ubiquitin hybrid chain pull-down assay proteomic screen (**Figure 3.2-7**). To determine whether the affinity observed between RAP80 and the poly-SUMO poly-ubiquitin hybrid chain type used in the proteomic screen was due to the tandem SIM UIMUIM of RAP80 a bacterially expressed RAP80 construct consisting of residues 1-150 which includes the tandem SIM UIMUIM domain (known as RAP80 SIM UIMUIM) was produced. This construct was then tested in the SUMO-ubiquitin hybrid chain pull-down assay to determine whether the SIM UIMUIM domain was

facilitating the affinity observed for RAP80 for the poly-SUMO poly-ubiquitin hybrid chain type used in the proteomic screen (**Figure 4.2-1, A**). As expected RAP80 SIM UIMUIM was found to be most abundant in the eluate from the poly-SUMO poly-ubiquitin containing pUb-4xSUMO2 resin type. Significantly less RAP80 SIM UIMUIM was detected in the eluates from K63 pUb, 4xSUMO2 and Ub-4xSUMO2 when compared to the eluate from the pUb-4xSUMO2 resin type (**Figure 4.2-1, B).** To determine whether RAP80 can interact with SUMO and ubiquitin *in vivo*, a GFP-RAP80 expressing U2OS cell line was obtained and tested to determine whether RAP80 and SUMO, or RAP80 and ubiquitin co-localised at the sites of DNA repair. The GFP-RAP80 expressing U2OS cells were treated with 4 Gys of IR and allowed to recover for 90 minutes prior to fixation with PFA and preparation for analysis. As expected, GFP-RAP80 is present in foci containing the DNA damage marker yH2Ax, after IR treatment (**Figure 4.2-1, C top panels**). Interestingly, ubiquitin (**Figure 4.2-1, C middle panels**) and SUMO (**Figure 4.2-1, C bottom panels**) are found to be present in GFP-RAP80 positive foci after IR treatment.

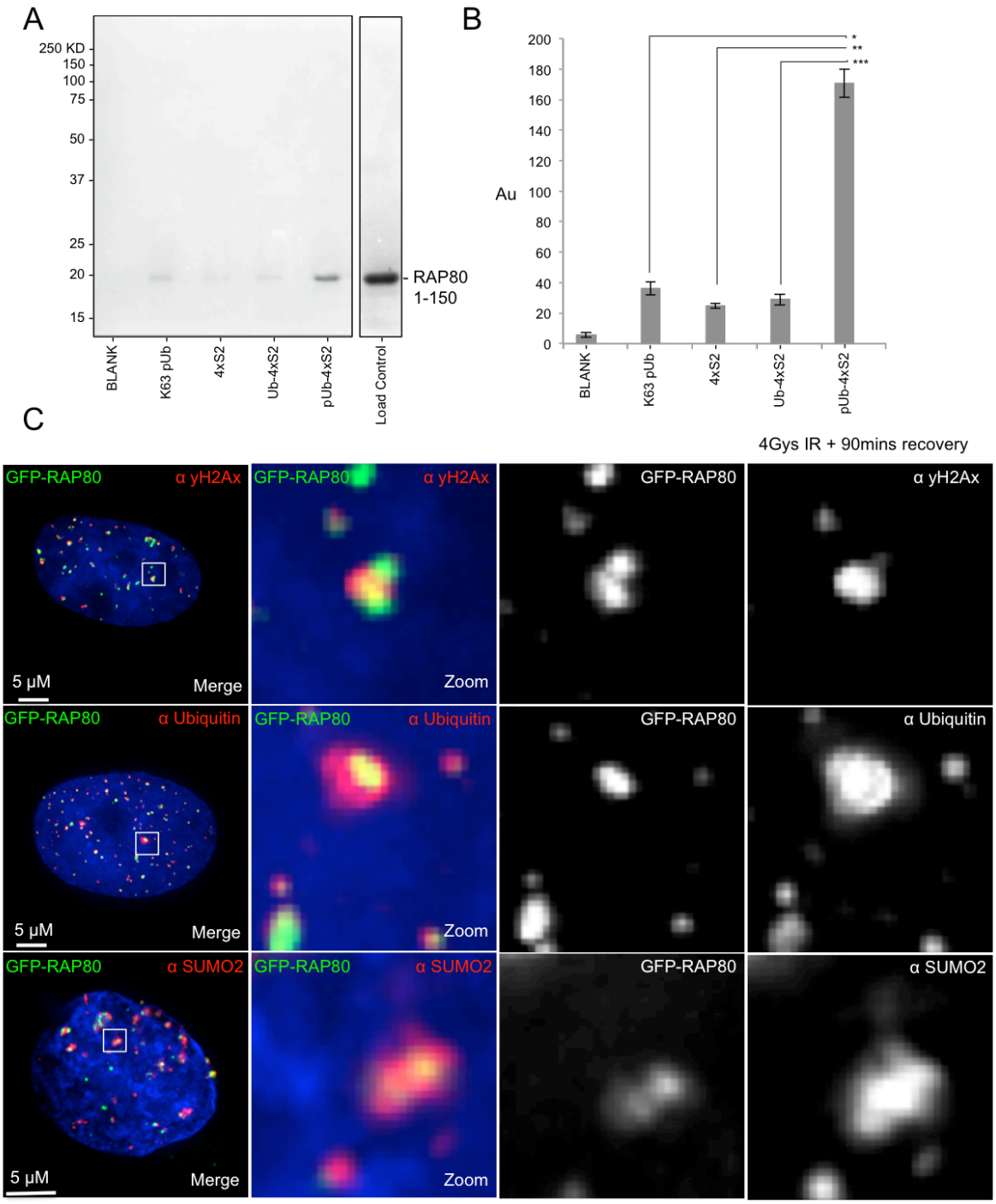


Figure 4.2-1 RAP80 shows affinity for SUMO and ubiquitin *in vitro* and *in vivo*.

A) SUMO-ubiquitin hybrid chain pull-down assay with recombinant RAP80 tandem SIM UIMUIM domain 1-150. Coomassie stained gel. The most prominent RAP80 band is associated with the pUb-4xSUMO2 resin type. Load Control represents 50% of total protein incubated with resins.

B) Quantification of SUMO-ubiquitin hybrid chain pull-down assay with RAP80 tandem SIM UIMUIM 1-150 domain. A significantly larger amount of RAP80 tandem SIM UIMUIM was found to be present in the eluate of the pUb-4xSUMO2 resin type than; Ub-4xSUMO2

(<0.005), 4xSUMO2 (<0.005), or K63 (<0.005) resin types. Experiment was conducted in triplicate. Error bars represent standard deviation from the mean.

C) Microscopy analysis of GFP-RAP80 (full length RAP80) U2OS cells after treatment with 4 Gys IR and recovery of 90 minutes. GFP-RAP80 is recruited to sites of yH2Ax foci after IR. Ubiquitin is associated with GFP-RAP80 foci after IR. SUMO is associated with GFP-RAP80 after IR. Images are single Z planes.

4.2.2 A proximity ligation assay identifies a close association by SUMO and ubiquitin in RAP80 positive foci after DNA damage

Establishing the presence of SUMO and ubiquitin separately in close association with GFP-RAP80 positive foci after IR treatment shows the potential context for RAP80 to interact with both SUMO and ubiquitin together *in vivo.*

To establish whether SUMO, ubiquitin, and RAP80 are all closely associated together *in vivo* required a new approach. The proximity ligation assay (PLA) uses antibodies coupled with short oligonucleotides, known as proximity probes, to establish whether two proteins are closely associated in a protein complex (**Figure 4.2-2**) (Gullberg et al., 2003).

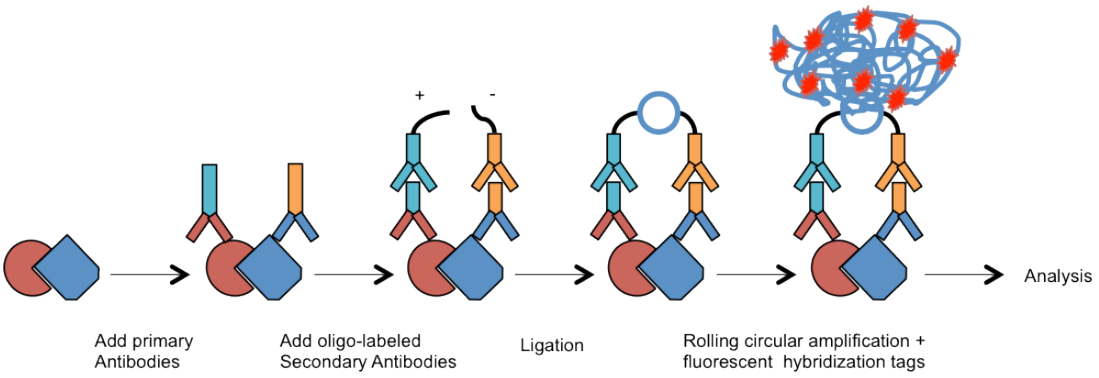


Figure 4.2-2 The Proximity Ligation Assay.

Schematic diagram represents the stages of a proximity ligation assay. Primary antibodies for two closely associated proteins are first added. Pairs of oligo-labeled secondary antibodies are then added, +/-, which correspond to the primary antibodies. The oligo are then ligated together, one + oligo to one – oligo. A rolling circular amplification is then carried out before fluorescent hybridization tags are added. Any closely associated proteins, within 40 nm, can then be visualized by microscopy.

The *in situ* PLA requires highly specific primary antibodies to firstly recognise two proteins that are in close proximity to one another. A second set of oligo-labeled secondary antibodies, specific for the primary antibodies are then added. If the two proteins are in close proximity, the addition of a short oligonucleotide linker ligates the oligo-labeled secondary antibodies together. A rolling circular amplification step is then carried out to create a large piece of single stranded DNA. Hybridisation tags specific for this newly generated DNA are then added and any positive PLA signals can be visualised by fluorescent microscopy (**Figure 4.2-2**). In this way only proteins that are closely associated, less than 40 nm, can generate a positive PLA signal. Commercially available specific antibodies have been raised against RAP80 (rabbit, Bethyl), SUMO (rabbit, Invitrogen; mouse, Hay lab generated), and ubiquitin (mouse, Merek), as such; the PLA is a potentially useful assay in the exploration of the association between RAP80, SUMO, and ubiquitin.

Firstly, PLAs were developed to probe the association of RAP80 with SUMO2, and RAP80 with ubiquitin (**Figure 4.2-3**). GFP-RAP80 expressing U2OS cells were treated with 4 Gys of IR and allowed to recover for 90 minutes. Cells were then fixed with PFA and prepared for PLA. Primary antibodies against RAP80 (rabbit) and either SUMO2 (mouse) or ubiquitin (mouse) were incubated with the fixed and permeabilised GFP-RAP80 U2OS cells. The control reactions contained the primary RAP80 antibody but neither SUMO2 or ubiquitin primary antibodies were present. Secondary oligo-labeled antibodies that had been raised against rabbit or mouse Immunoglobulin were then incubated with the cells. Ligation and amplification steps were carried out in accordance with **2.2.7.3**. When the SUMO2 antibody was omitted from the RAP80/SUMO2 PLA, no signal was detected from the assay (**Figure 4.2-3, A upper panels**). When both the RAP80 and SUMO2 antibodies were present, positive PLA signals were generated (**Figure 4.2-3, A lower panels**), suggesting that RAP80 and SUMO2 are in close proximity after DNA damage. The positive RAP80/SUMO2 PLA signals are shown to co-localize with the GFP-RAP80 foci (**Figure 4.2-3, B**). Warm colours represent complete co-localization between the PLA signal and GFP-

RAP80 foci. Cooler colours represent regions where the signals do not co-localize. As with the RAP80/SUMO2 PLA, when the ubiquitin antibody is omitted from the RAP80/ubiquitin PLA, no signal was detected from the assay (**Figure 4.2-3, C upper panels**). When both the RAP80 and ubiquitin antibodies were present, positive PLA signals were generated (**Figure 4.2-3, C lower panels**). This suggests again that RAP80 and ubiquitin are in close proximity after DNA damage. Several positive RAP80/ubiquitin PLA signals are shown to co-localize with the GFP-RAP80 foci (**Figure 4.2-3, D**). However, in both cases not all GFP-RAP80 positive foci are colocallised with positive PLA signals, perhaps suggesting diversity in the response.

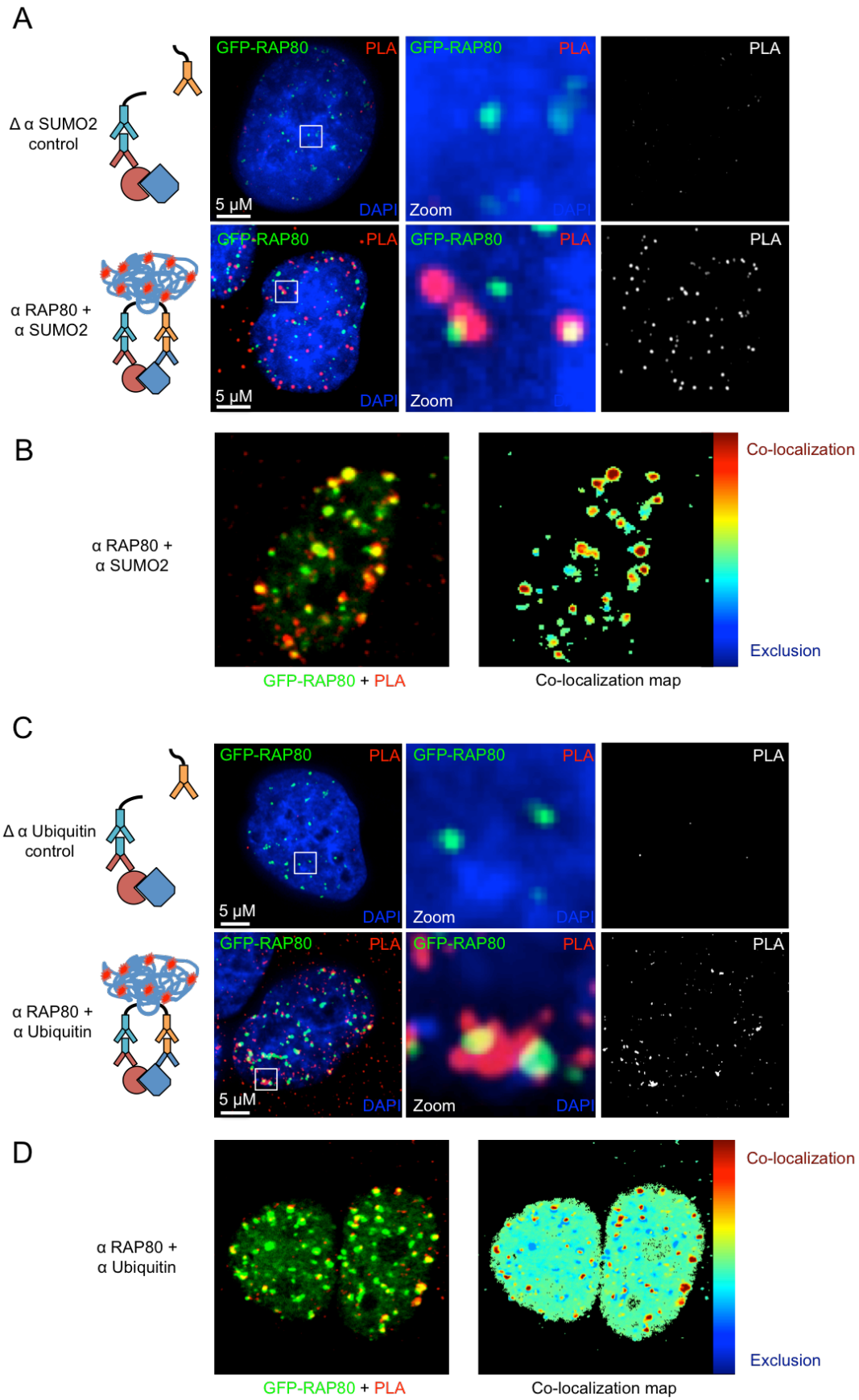


Figure 4.2-3 A Proximity Ligation Assay to detect SUMO and ubiquitin closely associated with RAP80.

A) Using the PLA assay to investigate the association of RAP80 and SUMO2 after 4Gys IR and 90 minute recovery. The top panel experiment omits the primary SUMO2 antibody for the PLA procedure as a negative control. Bottom panel experiment contains primary antibodies against SUMO2 and RAP80, thus the full compliment of antibodies required for a PLA signal to be generated. PLA signals are generated by this combination of antibody suggesting that SUMO2 and RAP80 are closely associated.
B) Quantification of the co-localization map of the PLA signal with GFP-RAP80 was performed using the normalized mean deviation product (nMDP) values ranging from -1 to1. Negative indexes are represented by cold colours (blue), exclusion, with the indexes greater than 0 represented by warm colours (red), co-localization. Images were taken using a spinning disk confocal microscope.
C) Using the PLA assay to investigate the association of RAP80 and ubiquitin after 4Gys IR and 90 minute recovery. Top panel experiment omits the primary ubiquitin antibody for the PLA procedure as a negative control. No PLA signal is detected. Bottom panel experiment contains primary antibodies against ubiquitin and RAP80, thus the full compliment of antibodies required for a PLA signal to be generated. PLA signals are generated by this combination of antibody suggesting that ubiquitin and RAP80 are closely associated.
D) Quantification of the co-localization map of the PLA signal with GFP-RAP80 was performed using the normalized mean deviation product (nMDP) values ranging from -1 to1. Negative indexes are represented by cold colours (blue), exclusion, with the indexes greater than 0 represented by warm colours (red), col-localization. Images were taken using a spinning disk confocal microscope.

Co-localization between the PLA signals generated by RAP80/SUMO2 and RAP80/ubiquitin pairs suggests that both SUMO and ubiquitin are not only present in RAP80 positive foci but are in close enough proximity to RAP80 that a direction interaction could be occurring.

A PLA was then established to detect a close association between SUMO2 and ubiquitin. SUMO2 (Rabbit) and ubiquitin (Mouse) antibodies were used as the primary PLA probes. Omitting the ubiquitin antibody from the SUMO2/ubiquitin PLA resulted in no positive signal being detected from the assay (**Figure 4.2-4, A upper panels**). When both the SUMO2 and ubiquitin antibodies were present, positive PLA signals were generated (**Figure 4.2-4, A lower panels**). Suggesting again that SUMO2 and ubiquitin are in close proximity after DNA damage inducing stimuli.

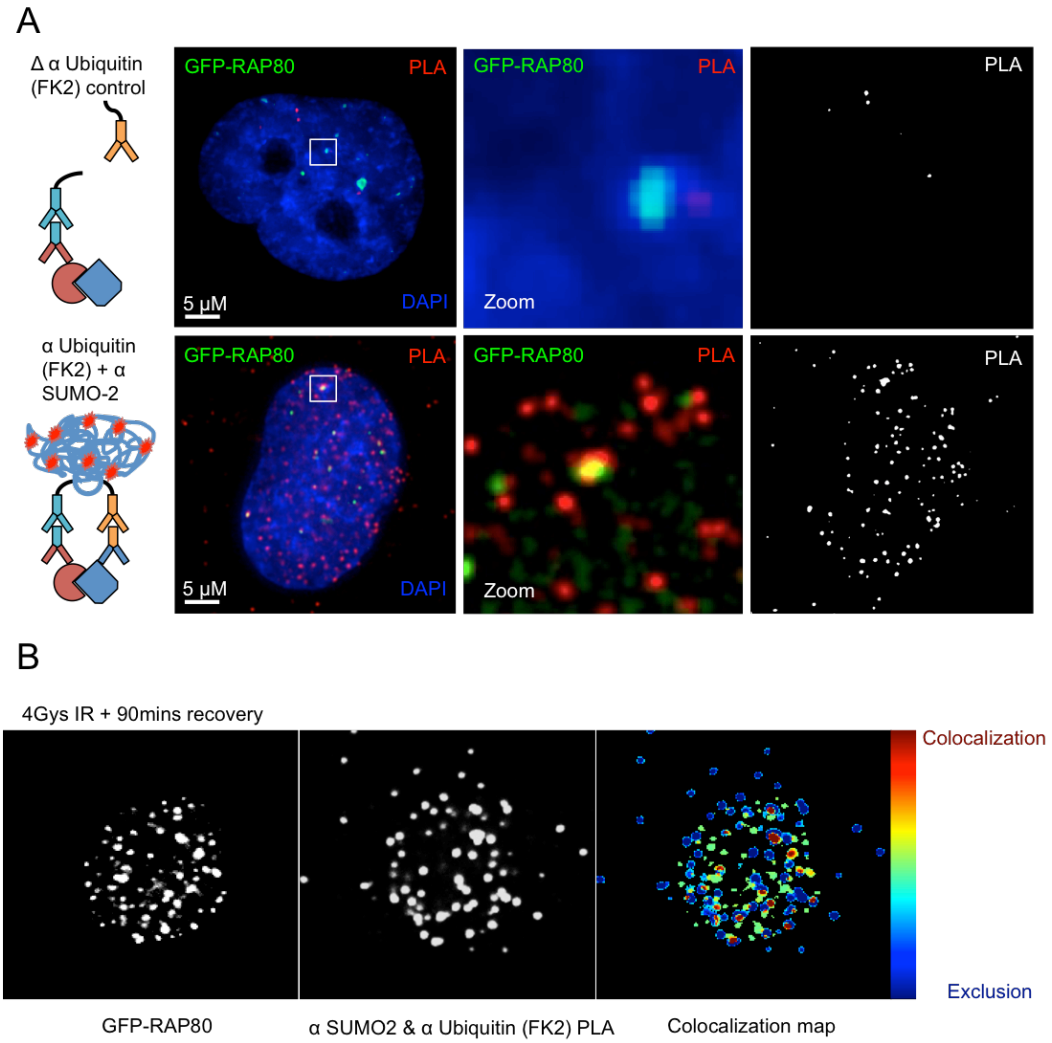


Figure 4.2-4 A Proximity Ligation Assay to detect SUMO-ubiquitin hybrid chains with RAP80 in close association.

A) Using the PLA assay to investigate the association of RAP80 with SUMO-ubiquitin hybrid chains after 4Gys IR and 90 minute recovery. Top panel experiment omits the primary ubiquitin antibody for the PLA procedure as a negative control. No PLA signal is detected. Lower panel experiment contains primary antibodies against SUMO2 and ubiquitin, thus the full compliment of antibodies required for a PLA signal to be generated. PLA signals are generated by this combination of antibody suggesting that SUMO2 and ubiquitin are closely associated.

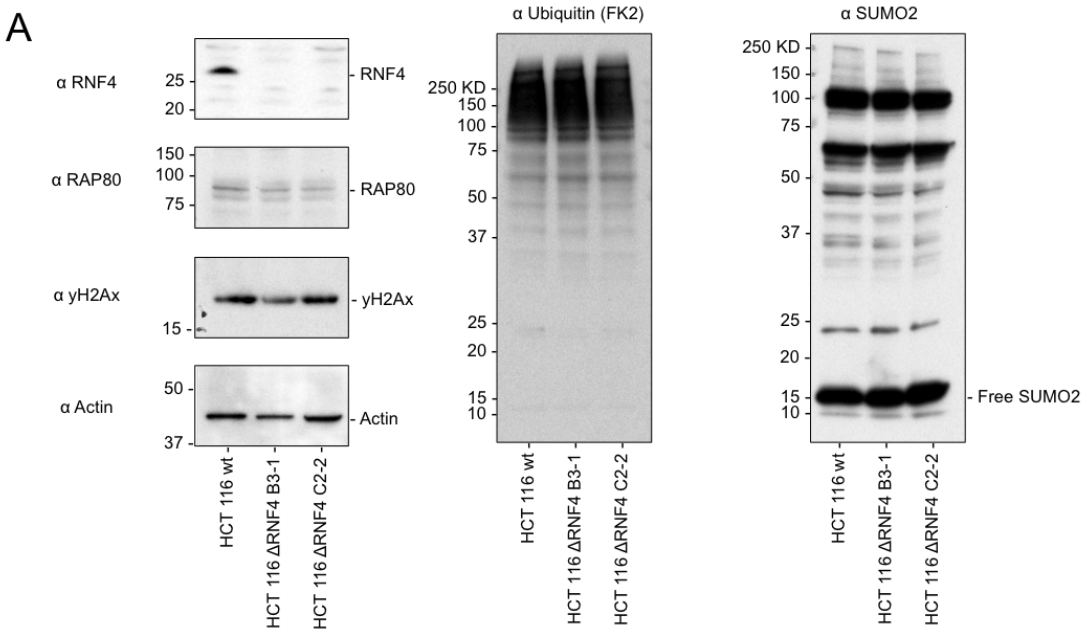
B) Quantification of the co-localization map of the PLA signal with GFP-RAP80 was performed using the normalized mean deviation product (nMDP) values ranging from -1 to1. Negative indexes are represented by cold colours (blue), exclusion, with the indexes greater than 0 represented by warm colours (red), col-localization. Images were taken using a spinning disk confocal microscope.

Interestingly, positive PLA signals generated by this combination of SUMO2/ubiquitin antibody were observed to co-localize with GFP-RAP80 foci after DNA damage (**Figure 4.2-4, B**). This suggests that RAP80, SUMO, and ubiquitin are indeed all closely associated in cells at the sites of DNA repair after DDR inducing stimulus. The close association seen by positive PLA signals generated by the SUMO2/ubiquitin PLA does not definitively suggest SUMO-ubiquitin hybrid chains, however, taken into account with the co-localisation of this signal with RAP80, a protein shown to have affinity for SUMO-ubiquitin hybrid chains, it does suggest that both SUMO and ubiquitin have a role with RAP80 at the sites of DNA repair.

4.2.2 The role of STUbLs in the recruitment RAP80 to sites of DNA repair

The SUMO targeted ubiquitin E3 ligase RNF4, has been shown to be present at the sites of DNA damage (Galanty et al., 2012; Luo et al., 2012; Yin et al., 2012). Recent work has described how knock out of RNF4 in DT40 chicken cells, lowers the levels of K63 linked poly-ubiquitin at the sites of DNA damage (Yin et al., 2012). Furthermore, RNF4 is suggested to have a role in recruiting BRCA1 to sites of DNA damage (Guzzo et al., 2012). Due to the potential role RNF4 could play in the recruitment of RAP80 to sites of DNA repair via the generation of SUMO-ubiquitin hybrid chains, a CRISPR/Cas9 generated knock out of RNF4 was generated in human colon carcinoma HCT116 cells.

Two ΔRNF4 HCT116 cell lines were gifted by Dr. J-F Maure, coded HCT116 ΔRNF4 B3-1, and HCT116 ΔRNF4 C2-2, along with the parental wild type HCT116 cell line. RNF4 was present in the parental wild type HCT116, but was absent in both HCT116 ΔRNF4 B3-1 and HCT116 ΔRNF4 C2-2 cell lines (**Figure 4.2-5, A top left hand panel**). RAP80, yH2Ax, ubiquitin, and SUMO protein levels, detectable via the antibodies used, were consistent between the wild type HCT116, HCT116 ΔRNF4 B3-1, and HCT116 ΔRNF4 C2-2 cell lines (**Figure 4.2-5, A**).


A
α RNF4
25
20
RNF4
α RAP80
150
100
75
RAP80
α yH2Ax
15
yH2Ax
α Actin
50
37
Actin
HCT 116 wt
HCT 116 ΔRNF4 B3-1
HCT 116 ΔRNF4 C2-2
α Ubiquitin (FK2)
250 KD
150
100
75
50
37
25
20
15
10
HCT 116 wt
HCT 116 ΔRNF4 B3-1
HCT 116 ΔRNF4 C2-2
α SUMO2
250 KD
150
100
75
50
37
25
20
15
10
Free SUMO2
HCT 116 wt
HCT 116 ΔRNF4 B3-1
HCT 116 ΔRNF4 C2-2

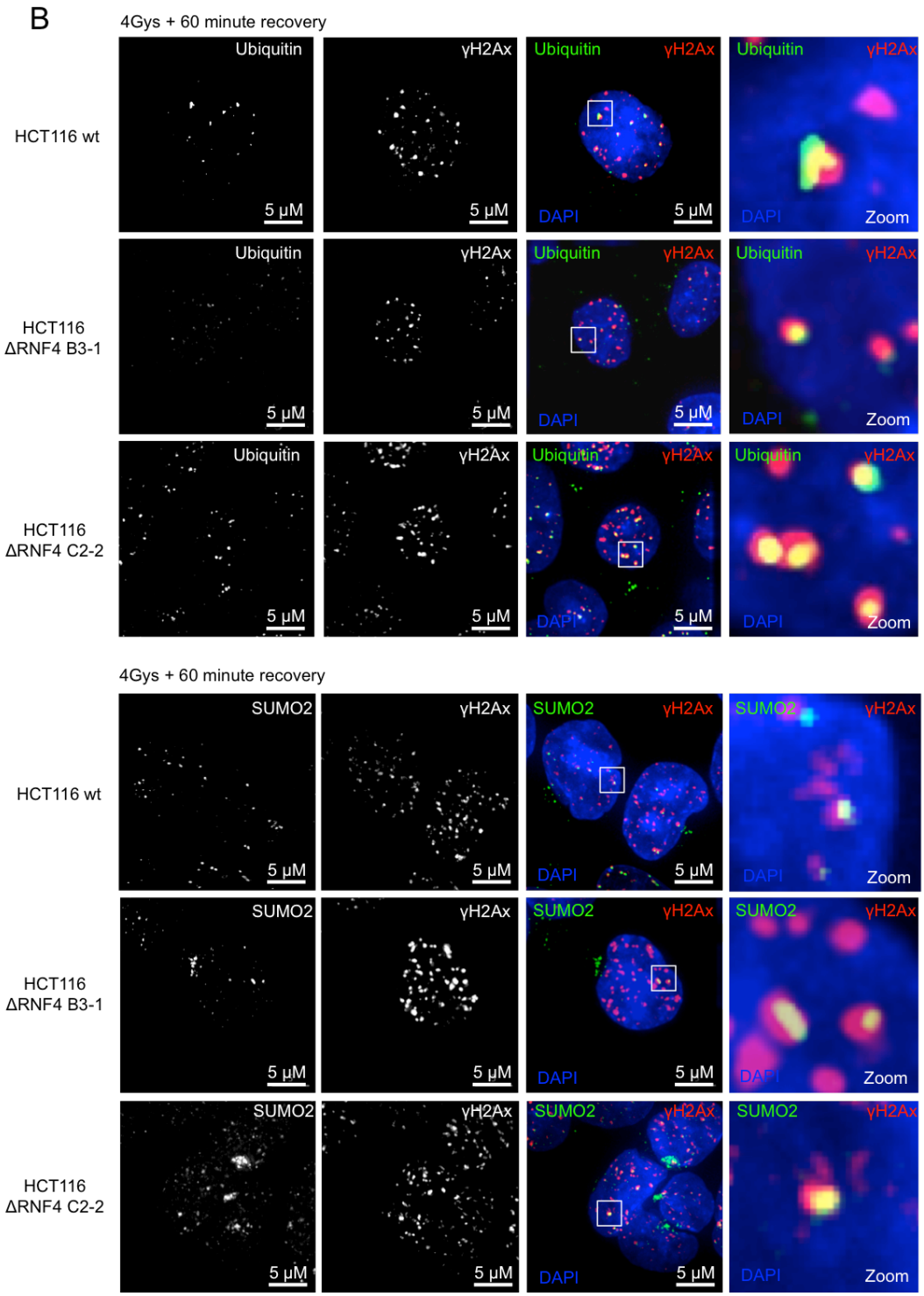


Figure 4.2-5 SUMO2 and ubiquitin are recruited to the sites of DNA damage in RNF4 knock out cells.

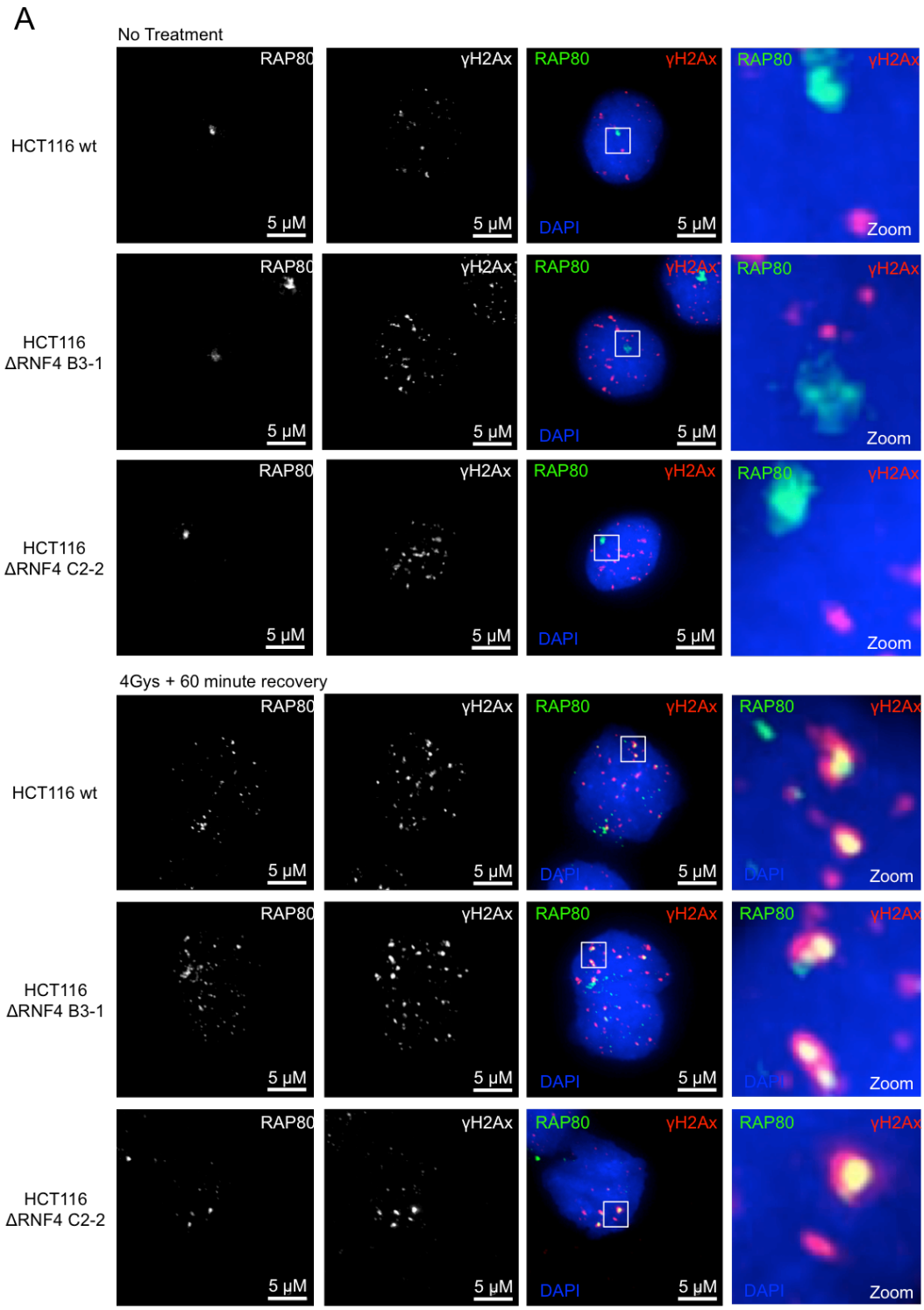
A) Western blot analysis of HCT116 wt, HCT116 ΔRNF4 clone B3-1, and HCT116 ΔRNF4 clone C2-2. RNF4 is present only in the extracts of the HCT116 wt cells. levels of RAP80, yH2Ax, Ubiquitin, and SUMO appear to be consistent between the HCT116 wt, HCT116 ΔRNF4 clone B3-1, and HCT116 ΔRNF4 clone C2-2 cell lines. Actin western blot is used as load control.

B) Immunofluorescence microscopy analysis of the HCT116 wt, HCT116 ΔRNF4 clone B3-1, and HCT116 ΔRNF4 clone C2-2 cell lines. HCT116 cell were checked for the presence of SUMO2 and ubiquitin at sites of DNA repair. After DNA damage was induced by 4 Gys of IR, SUMO2 and ubiquitin could be seen to be present at the site of DNA repair in even when RNF4 was knocked out.

Ubiquitin was present at the sites of DNA repair, shown via DNA damage marker, yH2Ax foci after IR in the wild type HCT116 cell line, and was also detected in the HCT116 ΔRNF4 B3-1 and HCT116 ΔRNF4 C2-2 cell lines (**Figure 4.2-5, B upper panels**). As with ubiquitin association, SUMO2 was present at the sites of DNA repair, in all three cell lines (**Figure 4.2-5, B lower panels**).

Although both SUMO and ubiquitin are present and seemingly unaffected by the lack of RNF4 at the site of DNA damage in the ΔRNF4 HCT116 cell lines this does not necessarily suggest that RAP80 will still be recruited as normal to sites of DNA damage. Thus, the percentage of RAP80 positive yH2Ax foci were compared in the wild type HCT116 cells and in the HCT116 ΔRNF4 B3-1 and C2-2 cell lines (**Figure 4.2-6**).

Levels of RAP80 recruitment to yH2Ax foci were then tested in both wild type HCT116 cells and the ΔRNF4 HCT116 cell lines. After IR treatment, RAP80 was observed to be present at the sites of DNA damage in wild type HCT116 cells after 60 minutes. Interestingly, RAP80 was also observed to be present at the sites of DNA damage in the RNF4 knock out cell line (**Figure 4.2-6, A**). This suggests that RNF4 is not required for RAP80 recruitment to the sites of DNA damage. The observed number of cells with RAP80 positive yH2Ax foci in wild type HCT116 was not significantly altered in either the HCT116 ΔRNF4 B3-1 or HCT116 ΔRNF4 C2-2 cell lines (**Figure 4.2-6, B**). RAP80 recruitment to sites of DNA damage in HCT116 cells was therefore not ablated by the complete knockout of RNF4.


A
No Treatment
HCT116 wt
HCT116 ΔRNF4 B3-1
HCT116 ΔRNF4 C2-2
4Gys + 60 minute recovery
HCT116 wt
HCT116 ΔRNF4 B3-1
HCT116 ΔRNF4 C2-2
RAP80
γH2Ax
DAPI
5 μM
Zoom

B

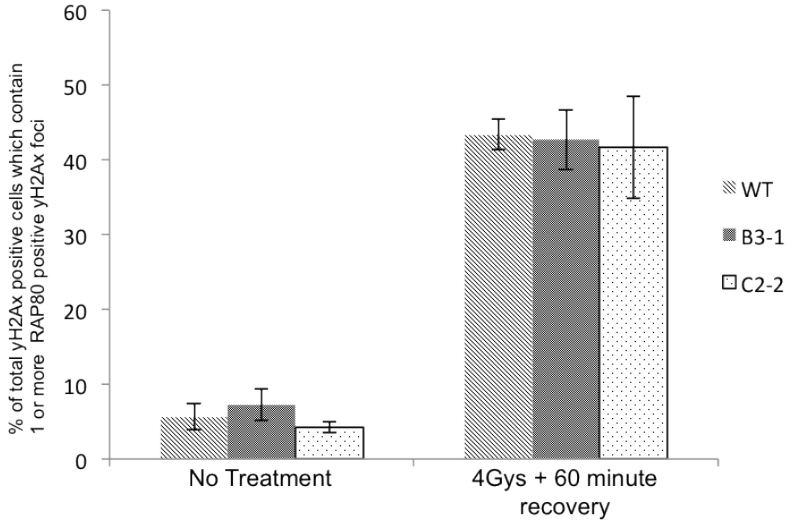


Figure 4.2-6 RAP80 is recruited to the sites of DNA damage in RNF4 knock out cells.

A) Immunofluorescence microscopy analysis of RAP80 at the sites of DNA repair. HCT116 cell were checked for the presence of RAP80 at sites of DNA repair. After DNA damage was induced by 4 Gys of IR, RAP80 could be seen to be present at the site of DNA repair even when RNF4 was knocked out.

B) Quantification of co-locaization between endogenous RAP80 and yH2Ax in wild type HCT116, HCT116 ΔRNF4 B3-1, and HCT116 ΔRNF4 C2-2 cell lines. No significant difference was observed in percentage of cells containing both RAP80 and yH2Ax foci and possessed 1 or more RAP80 positive yH2Ax foci from 3 independent experiments n<80 cells counted. Error bars represent standard deviation from the mean.

Along with RNF4 another STUbL, RNF111, has been described to be present at the sites of DNA damage. RNF111 contains two high fidelity SIM domains in its N-terminal region (Sun & Hunter 2012). These SIM domains have been shown, like RNF4, to be important in the recruitment and E3 activity of the protein (Poulsen et al., 2013). Since knocking out RNF4 had no observable effect on the recruitment of RAP80 to the sites of DNA damage, RNF111 was also knocked down via siRNA to test for a possible role of this STUbL in the recruitment of RAP80 to the site of DNA damage.

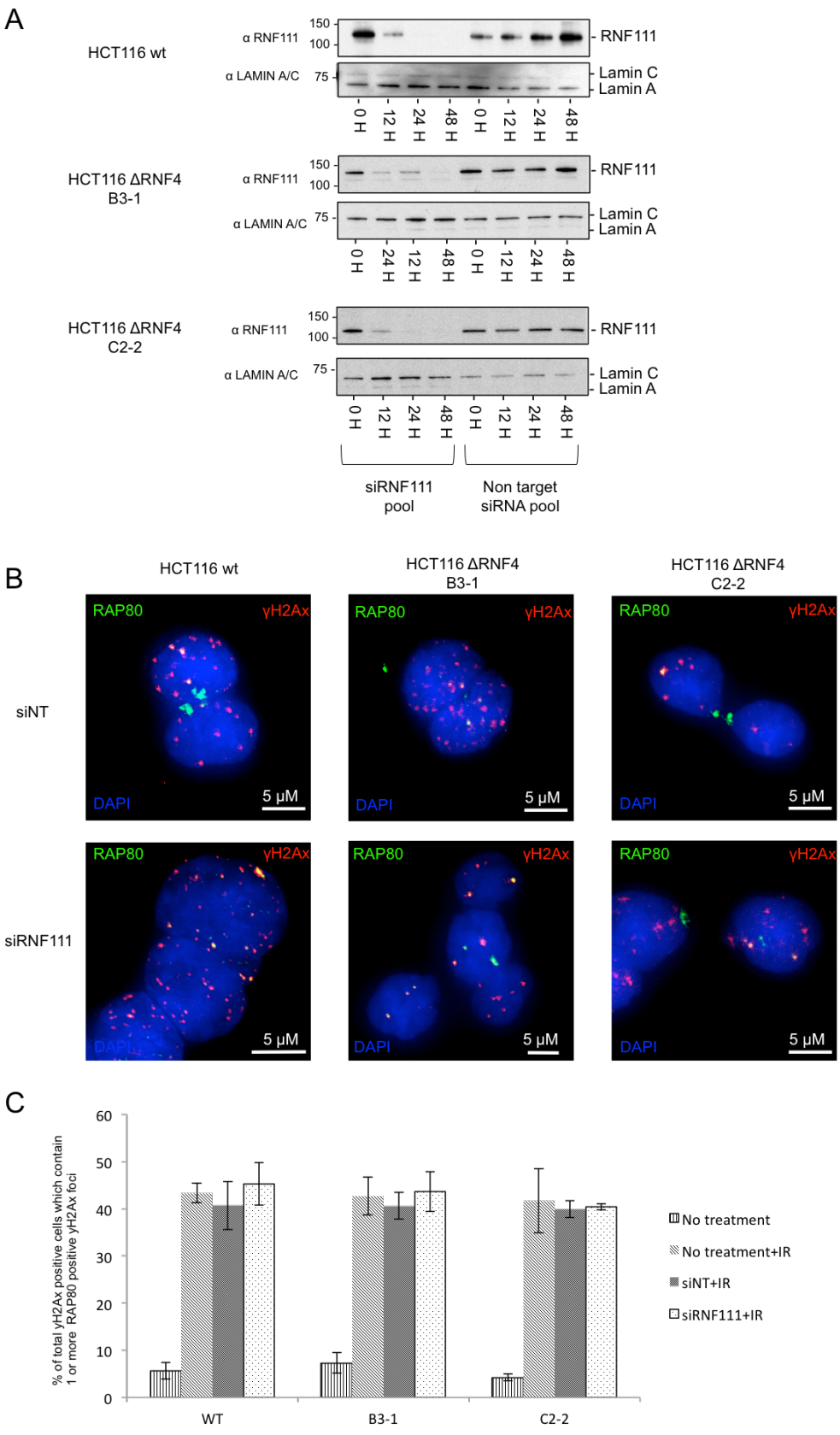


Figure 4.2-7 Knock down of the STUbL RNF111 does not disrupt RAP80 recruitment to sites of DNA repair, even in a RNF4 null background.

A) Immunofluorescence staining microscopy analysis of endogenous RAP80 recruitment to yH2Ax foci under RNF111 knockdown on wild type HCT116, HCT116 ΔRNF4 B3-1, HCT116 ΔRNF4 C2-2 cell lines. RAP80 is still recruited to site of DNA damage after siRNF111. Analyzed after 4 Gys of IR and 60 minute recovery.
B) Western blot analysis of RNF111 siRNA knockdown. After 62 hours RNF111 was efficiently knocked down by a pool of siRNF111 in wild type HCT116, HCT116 ΔRNF4 B3-1, HCT116 ΔRNF4 C2-2 cell lines.
C) Quantification of RAP80 positive foci in ΔRNF4 siRNF111 HCT116 cell lines. No significant difference was observed in percentage of cells containing both RAP80 and yH2Ax foci and possessed 1 or more RAP80 positive yH2Ax foci. Quantification from 3 independent experiments n<80 cells counted. Error bars represent standard deviation from the mean.

RNF111 was effectively knocked down after 62 hours of treatment with siRNA for RNF111, While siNT had no effect on the levels of RNF111 in the wild type HCT116, HCT116 ΔRNF4 B3-1, or the HCT116 ΔRNF4 C2-2 cell lines (**Figure 4.2-7, A**). RAP80 was recruited to the sites of DNA repair in the wild type HCT116 cells after RNF111 siRNA treatment (**Figure 4.2-7, B lower left panel**), and again RAP80 was recruited to yH2Ax positive foci both the ΔRNF4 HCT116 cell lines after siRNA treatment for RNF111 (**Figure 4.2-7, B lower middle and right panels**). No observable divergence was seen in percentage of cells reporting RAP80 positive yH2Ax foci after RNF111 siRNA treatment in either wild type HCT116, HCT116 ΔRNF4 B3-1, or HCT116 ΔRNF4 C2-2 (**Figure 4.2-7, C**).

Another notable observation made when studying the HCT116 ΔRNF4 cell lines was the large number of yH2Ax positive foci that were present in cells under normal cell culture growth conditions. Non-treated wild type HCT116 cell contained very low number of yH2Ax foci, in most cases less than 5. However, both ΔRNF4 cell lines contained markedly more yH2Ax foci under normal growth conditions (**Figure 4.2-8, A & B**). Over 50% of the HCT116 ΔRNF4 cells that were positive for yH2Ax contained more than 5 yH2Ax foci. As both the HCT116 ΔRNF4 B3-1 and HCT116 ΔRNF4 C2-2 cell lines appeared to double and grow at a similar rate to the wild type cells, this appeared to suggest that the cells containing RNF4 could carry damage more easily through the cell cycle. Each of the three HCT116 cell lines was then

prepared for flow cytometry and cell cycle analysis to determine if either the HCT116 ΔRNF4 B3-1 or HCT116 ΔRNF4 C2-2 cell lines differed to the wild type HCT116 cells in cell cycle composition.

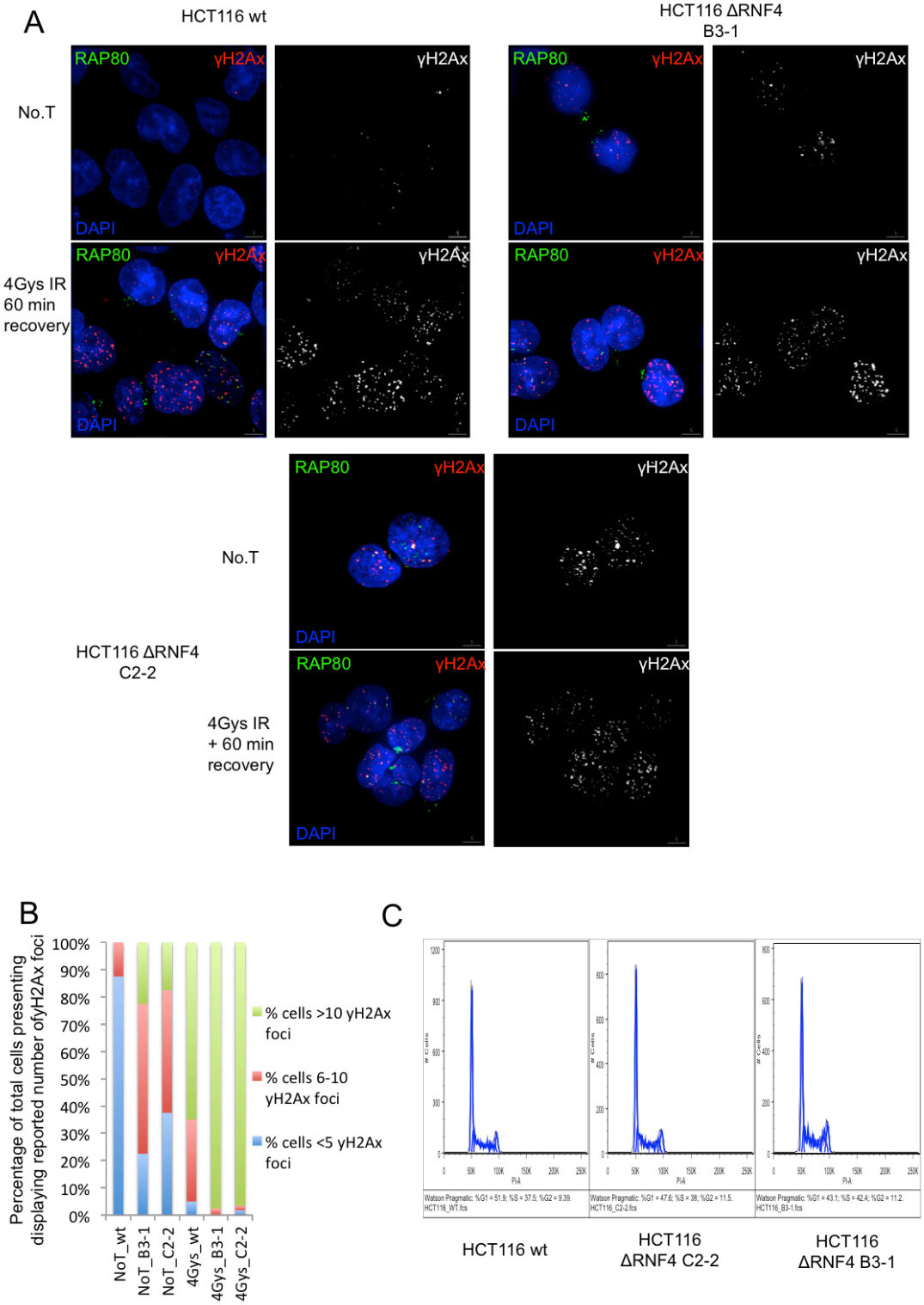


Figure 4.2-8 RNF4 knock out HCT116 cell line show higher levels of endogenous γH2Ax foci than wild type HCT116 cells.

A) Immunofluorescence microscopy analysis of yH2Ax foci number in wild type HCT116, HCT116 ΔRNF4 B3-1, and HCT116 ΔRNF4 C2-2 cells. Under endogenous cell culture conditions both HCT116 ΔRNF4 cell lines present with a higher number of yH2Ax foci per cell than the wild type HCT116 cell line. All three cell lines show an increase in yH2Ax foci formation after DNA damage.
B) Quantitation of yH2Ax foci numbers present in each cell line. Greater than 50% of HCT116 ΔRNF4 B3-1 and HCT116 ΔRNF4 C2-2 contained greater than 5 yH2Ax foci under endogenous conditions, while 87% of wild type HCT116 cells contained less than 5 yH2Ax foci. After DNA damage inducing stimuli the number of yH2Ax foci increased in all three cell lines. Results are the average of 4 independent experiments, n>40 cells.
C) Cell cycle analysis of wild type HCT116, HCT116 ΔRNF4 B3-1, HCT116 ΔRNF4 C2-2 cell lines. In each cell type the majority of cells were observed to be in the G1 cell cycle stage. No discernable differences were noted between the wild type HCT116 cells and the ΔRNF4 HCT116 cell lines.

However, no discernable differences were observed between the wild type and ΔRNF4 HCT116 cell lines under endogenous cell culture growth conditions (**Figure 4.2-8, C**). The majority of cells from each cell line were in G1 and no large populations of, G0 or S phase cells indicating cells that may be facing difficulty progressing through the cell cycle were observed. Thus although the HCT116 cells lacking RNF4 seemingly contained more yH2Ax foci under normal cell culture growth condition when compared to the parental wild type HCT116 cells this did not disrupt cell cycle progression.

4.2.3 Exploring the effect of SUMO proteases on RAP80 associated proteins

Understanding the interplay between RAP80, SUMO and ubiquitin will help to understand whether SUMO-ubiquitin hybrid chains play a role in recruiting RAP80 to the sites of DNA damage. SUMO proteases are a powerful tool to determine any interplay between RAP80 and SUMOylated proteins.

Previous work has suggested that SUMO-ubiquitin hybrid chains anchor RAP80 near to the sites of DNA damage (Guzzo et al., 2012), and earlier work in this thesis supports the view that RAP80, SUMO and ubiquitin

are in close association at the sites of DNA damage. If RAP80 is indeed anchored to the sites of DNA damage then removing the SUMO moiety of the SUMO-ubiquitin hybrid chain hypothetically anchoring RAP80 would separate the proteins anchoring the hybrid chain from RAP80, thus acting to remove the anchoring protein from any RAP80 containing protein complex (**Figure4.2-9, A**). In this manner an experiment was devised utilizing a SENP1 protease treatment of HEK 293 nuclear cell extracts in conjunction with a RAP80 immunoprecipitions (IPs) to evaluate the association of SUMO, ubiquitin, and RAP80 associated protein complexes.

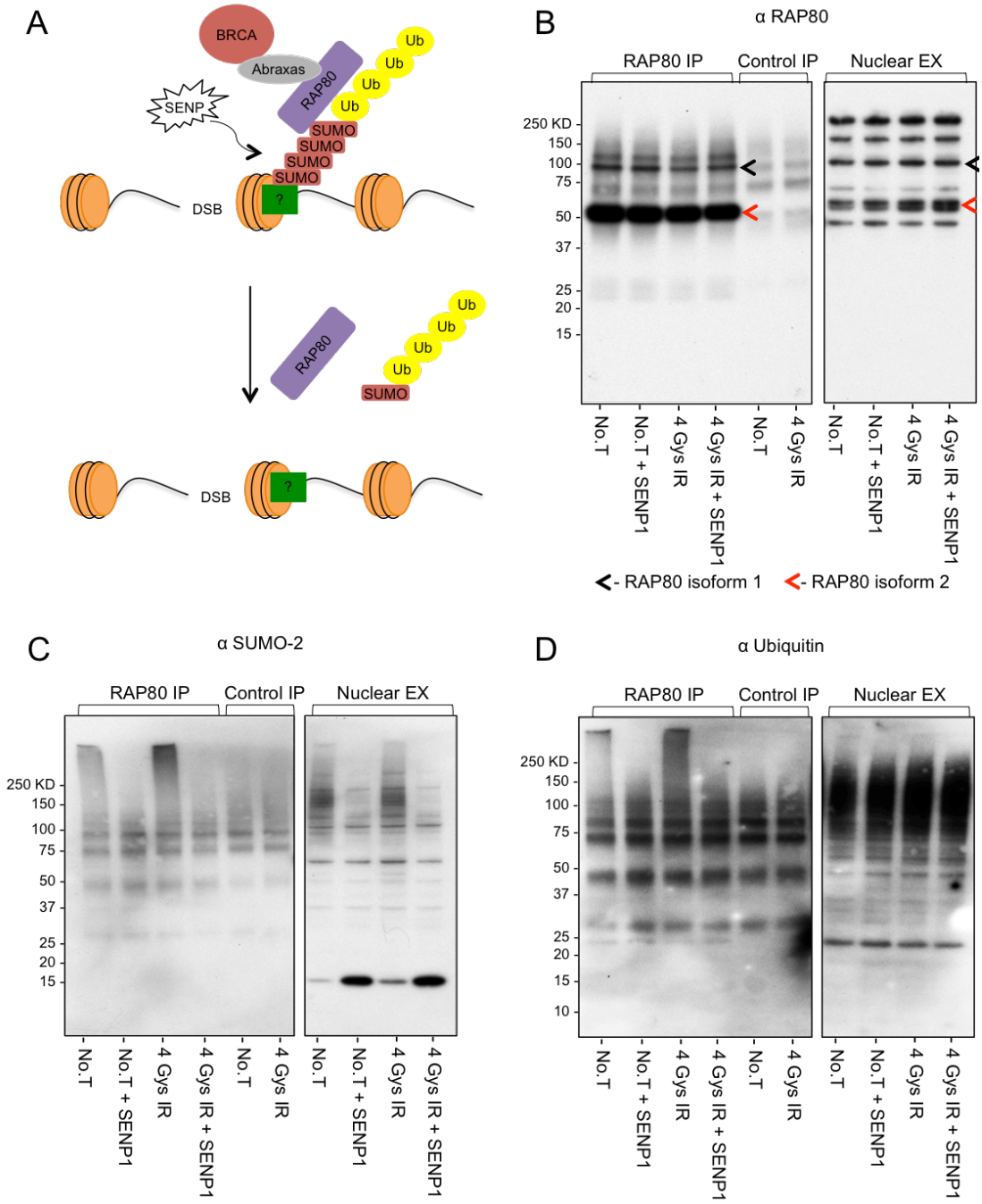


Figure 4.2-9 Hunting for SUMO-ubiquitin hybrid chain anchoring proteins using SENP1 protease treatment.

A) Schematic diagram represents the hypothesized action of a SUMO protease on a hybrid chain of SUMO and ubiquitin anchoring RAP80 to the sites of DNA damage.

B) C) D) Western blot analysis of RAP80 immunoprecipitations from HEK 293 nuclear cell extracts that have been treated with SENP1 shows a close association of SUMO and ubiquitin in regards to the material immunoprecipitated with RAP80. Cells were either left untreated (No.T) or irradiated with 4 Gys and left for 60 mins to recover prior to nuclear extracts being prepared. extracts were treated with 10 μM SENP1 for 4 hours at room temperature, control pull downs were left at room temperature for 4 hours. B) α RAP80

western blot shows RAP80 is immunoprecipitated only after the RAP80 IP. Arrow heads indicate RAP80; isoform 1 is denoted by black arrow heads and isoform 2 is denoted by red arrow heads C) α SUMO2 western blot shows decrease of high molecular weight SUMO conjugates associated with RAP80 immunoprecipitated material after SENP1 treatment. This decrease is also present in the nuclear extracts after SENP1 treatment D) α ubiquitin western blot shows a decrease in the high molecular weight ubiquitin conjugates associated with the RAP80 immunoprecipitated material after SENP1 treatment. However, the same decrease in high molecular weight ubiquitin conjugates is not seen in the HEK 293 nuclear extracts.

HEK293 cells were grown until 100% confluent on 20 10 cm dishes. 10 10 cm dishes of 100% confluent cells were treated with 4 Gys IR followed by 60 minutes recovery period, and the remaining 10 dishes were left untreated. Nuclear extracts were then prepared from both pools of cells. 2.5 mg of the cell extracts from the IR treated cells, and 2.5 mg from the untreated pool of cells were then incubated separately with SENP1 in accordance with **2.2.8.3**. This resulted in four nuclear extract conditions; No treatment (No.T), No treatment with SENP1 treatement (No.T+SENP1), IR treatment (IR), IR treatment with SENP1 treatment (IR+SENP1). Immunoprecipitations with the RAP80 antibody were then performed, before SDS-PAGE and western blot analysis were carried out. RAP80 was successfully immunoprecipated from all cell extract treatment conditions via this protocol (**Figure 4.2-9, B**). The SENP1 treatment was also successful in reducing the SUMOylation landscape of the nuclear cell extracts. Material associated with the RAP80 IPs from the non-SENP1 treated nuclear extracts, No.T and IR, contained highly SUMO2-modified material, where as the material associated with the RAP80 IPs associated with the SENP1 treated extracts contained less SUMOylated material (**Figure 4.2-9, C**). This again suggests that RAP80 associated protein complexes contain SUMO2. As expected, SENP1 treatment had no observable effect on the global ubiquitylation landscape of the No.T or IR treated nuclear extracts (**Figure 4.2-9, D**). Importantly however, the SENP1 treatment reduced the apparent high molecular weight ubiquitin conjugates associated with the RAP80 immunopreciptated material. This suggests that the ubiquitin and SUMO modifications associated with RAP80 are highly cooperative, in that a portion of the proteins that are

associated with RAP80 in a SUMO dependent manner are also modified with ubiquitin.

Mass spectrometry analysis was then conducted to determine the effect SENP1 treatment had on RAP80 immunoprecipitated material from untreated and IR treated cell extracts.

The RAP80 IP experiment detailed above was repeated in triplicate before peptides were created in accordance with **2.2.8.1**. Samples were then processed using MaxQuant. Analysis of the mass spectrometry data then showed that as expected, RAP80 was consistently immunoprecipated from all cell extract conditions (No.T, No.T+SENP1, IR, and IR+SENP1) (**Figure 4.2-10**). As expected, levels of SUMO1 and SUMO2 detected from the RAP80 IPs were lower from the cell extracts that had been treated with SENP1 when compared to the extracts that had not been treated with SENP1. Interestingly as seen in the western blot analysis, SENP1 treatment lowered the intensity of ubiquitin detected when compared to the RAP80 IPs from cell extracts not treated with SENP1, suggesting that the relative abundance of ubiquitin was lower after SENP1 treatment. This again suggests that material associated with RAP80 in a SUMO specific manner is associated with ubiquitin, as removing the SUMOylated material associated with RAP80 also reduced the relative abundance of ubiquitin detected by mass spectrometry (**Figure 4.2-10**).

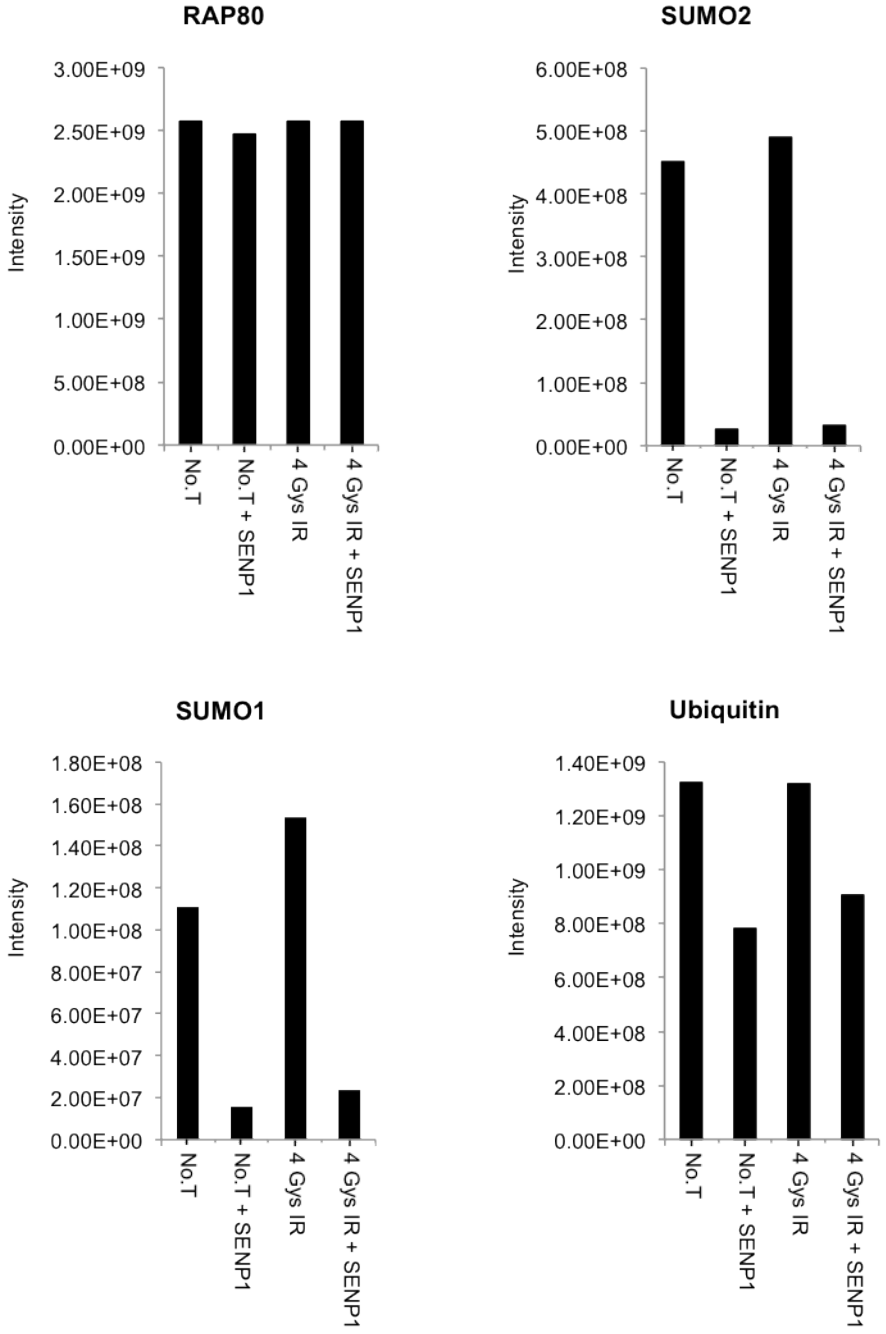
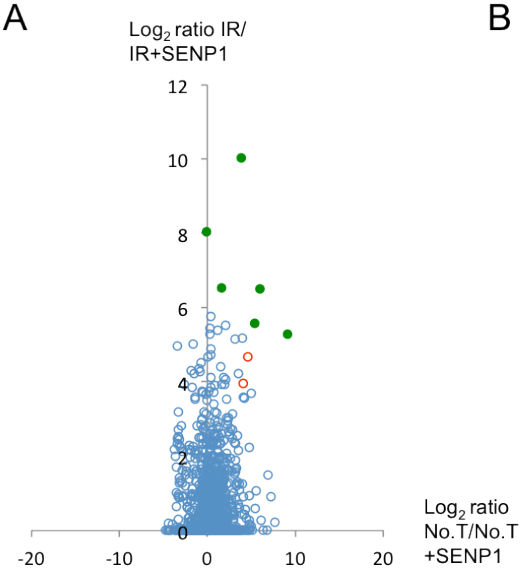


Figure 4.2-10 Normalised intensity charts for RAP80, SUMO2, SUMO1 and ubiquitin for RAP80 IPs from HEK 293 nuclear cell extracts.

Intensity data was normalised against the intensities reported for the 10 most abundantly found proteins across all IPs. RAP80 was consistently immunoprecipitated by the RAP80 antibody. Relative abundance of both SUMO2 and SUMO1 was severely reduced from RAP80 IPs from SENP1 treated HEK293 nuclear cell extracts. Relative abundance of ubiquitin was reduced from RAP80 IPs from SENP1 treated HEK293 nuclear cell extract.

Global analysis of the data was then carried out to determine if the abundance of any proteins could be observed that decreased specifically with respect to the SENP1 treatment prior to the immunoprecipitation of RAP80 (**Figure 4.2-11, A**). Three of the most significantly reduced proteins from SENP1 treated RAP80 IPs in both non-treated and damaged cell extracts were C5orf25, TOPORs, and RAD54L2 (**Figure 4.2-11, B & C**). All three proteins are known to be SUMO modified (Tammsalu et al 2014) (**Figure 4.2-11, B**). Thus it is possible that these proteins are associated with RAP80 when SUMOylated. C5orf25 has a single SUMOylation site, TOPORs has four known SUMOylation sites, and RAD54L2 has 17 known SUMOylation sites. C5orf25 is a known SUMO interacting protein (Hunter & Sun 2012), but has also been observed like RAP80, to have affinity for SUMO-ubiquitin hybrid chains (**Figure 3.2-10**). TOPORs, a ubiquitin E3 ligase, has been shown to interact with SUMO and has also been observed to show SUMO E3 ligase activity (Wenger et al., 2005). RAD54L2, a DNA helicase, has been shown to contain SIMs and interact with SUMO (Ogawa et al., 2009). All three proteins have the potential then to not only interact with SUMO non-covalently but also to be SUMOylated. Thus, these proteins can only be said to interact with RAP80 associated material in a SUMO specific manner as these proteins may be interacting non-covalently with other SUMOylated material associated with RAP80. However, the known SUMOylation sites located within C5orf25, TOPORs, and RAD54L2 allows for the potential of these proteins to be modified by SUMO-ubiquitin hybrid chains.

A



B

Protein Name	Protein Identifier	SUMO modified
Uncharacterized protein C5orf25	C5orf25	✓ K92
39S ribosomal protein L14, mitochondrial	MRPL14	
DmX-like protein 1	DMXL1	
E3 ubiquitin-protein ligase Topors	TOPORS	✓✓✓✓ K11, K734, K754, K885
Helicase ARIP4	RAD54L2	✓x17
Thioredoxin	PDIA3	✓ K366

C

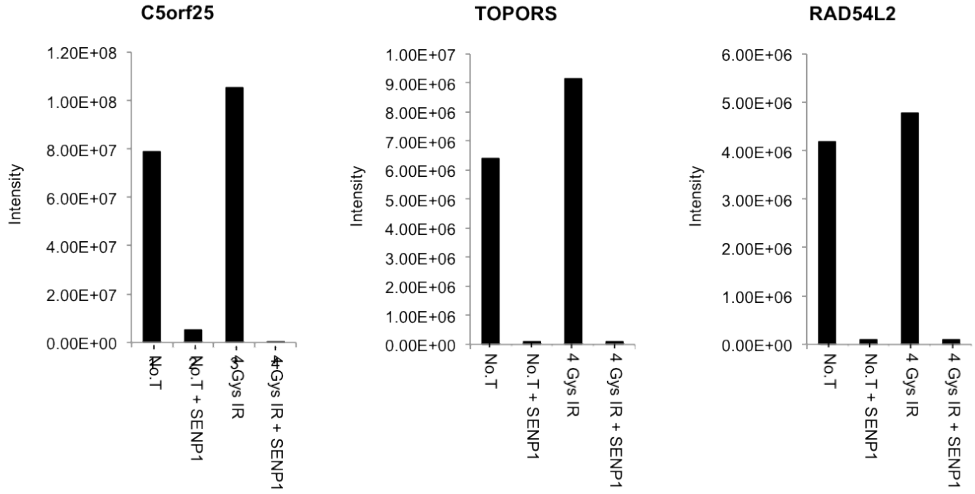


Figure 4.2-11 Analysis of SENP1 treatment on RAP80 associated proteins.

A) Red outlined points are SUMO2 and SUMO1. Green points are proteins that are significantly less abundant in the RAP80 IPs from the SENP1 treated cell extracts when compared to the RAP80 IPs from untreated cell extracts.

B) List of top 5 most changed proteins after SENP1 treatment of RAP80 IPs.

C) Normalised intensity charts for C5orf25, TOPORs, and RAD54L2. The relative abundance of C5orf25, TOPORs, and RAD54L2 were all severely reduced from RAP80 IPs from SENP1 treated HEK293 nuclear cell extracts when compared to RAP80 IPs from untreated HEK293 nuclear cell extracts.

The interplay between SUMO and ubiquitin in the context of RAP80 is quite complex. Decreasing the SUMOylation landscape of HEK 293 nuclear

cell extracts prior to RAP80 immunoprecipitiation severely reduces the RAP80 associated hyper-ubiquitylated material, suggesting that SUMOylation is key for the association of a portion of ubiquitylated material with RAP80. It is interesting to note that C5orf25, TOPORs and RAD54L2 are all known to be SUMO modified. This suggests that these proteins could be interacting with RAP80 when SUMO modified, which in turn allows for the possibility that SUMO-ubiquitin hybrid chains could form on these proteins. This approach identifies SUMO modification of RAP80 associated proteins as a key factor in the recruitment of ubiquitin modifications to RAP80 but to identify whether this cooperation between SUMOylation and ubiquitylation is the result of SUMO-ubiquitin hybrid chains requires a different approach.

4.2.4 A SILAC based RAP80 IP gel shift assay utilizing USP2 and SENP1 to evaluate the potential for SUMO-ubiquitin hybrid chain anchoring proteins

Treating cell extracts with SENP1 prior to RAP80 immunoprecipitation has shown that some ubiquitin-associated proteins appear to be linked with RAP80 associated material in a SUMO dependent manner. The intensity of ubiquitin recorded from the RAP80 associated material was reduced from SENP1 treated cell extracts when compared to the intensity of ubiquitin reported form cell extracts that had not been treated with SENP1. This suggested a high degree of cooperativity between SUMO and ubiquitin modifications associated with RAP80, however, this does not definitively answer whether RAP80 is indeed interacting with proteins modified with SUMO-ubiquitin hybrid chains. Thus, an experiment was designed to determine whether or not this apparent high degree of cooperativity was signalling the presence of SUMO-ubiquitin hybrid chain modified proteins or just proteins that are modified with homotypic chains of both SUMO and ubiquitin (**Figure 4.2-12, A**). IR treated Light, Medium, and Heavy SILAC labelled HEK 293 cell extracts were prepared before separate RAP80 IPs were performed on each of the SILAC labelled nuclear extracts. Protease treatments were then carried out on the RAP80 IP associated material; the Light RAP80 IP would be left untreated as a control, the Medium RAP80 IP would be treated with USP2 CaD to remove any ubiquitin modifications, and the Heavy RAP80 IP would be treated with SENP1 to remove any SUMO

modifications. In the context of SUMO-ubiquitin hybrid chains this would allow the Light RAP80 IP sample to contain un-altered hybrid chains, but both the Medium RAP80 IP and Heavy RAP80 IP would no longer contain full SUMO-ubiquitin hybrid chains. In theory, the Medium RAP80 IP would now contain proteins only modified with the SUMO element of the SUMO-ubiquitin hybrid chain and the Heavy RAP80 IP having the hybrid chains removed completely. The RAP80 IPs from the three SILAC conditions could then be combined, and separated by molecular weight via SDS-PAGE. The resulting protein-containing gel could then be cut into multiple slices and protein intensities could be compared by apparent molecular weight (or slice) and SILAC label for each protein. If the protein was modified by a SUMO-ubiquitin hybrid chain then the Light labelled protein would be observed to be a higher molecular weight than the predicted molecular weight of the unmodified protein, nearer the top of the gel. The Medium labelled protein still containing the SUMO portion of the hybrid chain would be observed at a higher molecular weight than the predicted molecular weight of the unmodified protein but would not be observed as near to the top of the gel as the Light labelled sample. The Heavy labelled sample would be close to the expected size of the unmodified protein and lower in the gel when compared to the Medium and Light RAP80 IP samples, as the SUMO-ubiquitin hybrid chains would be removed (**Figure 4.2-12, A**).

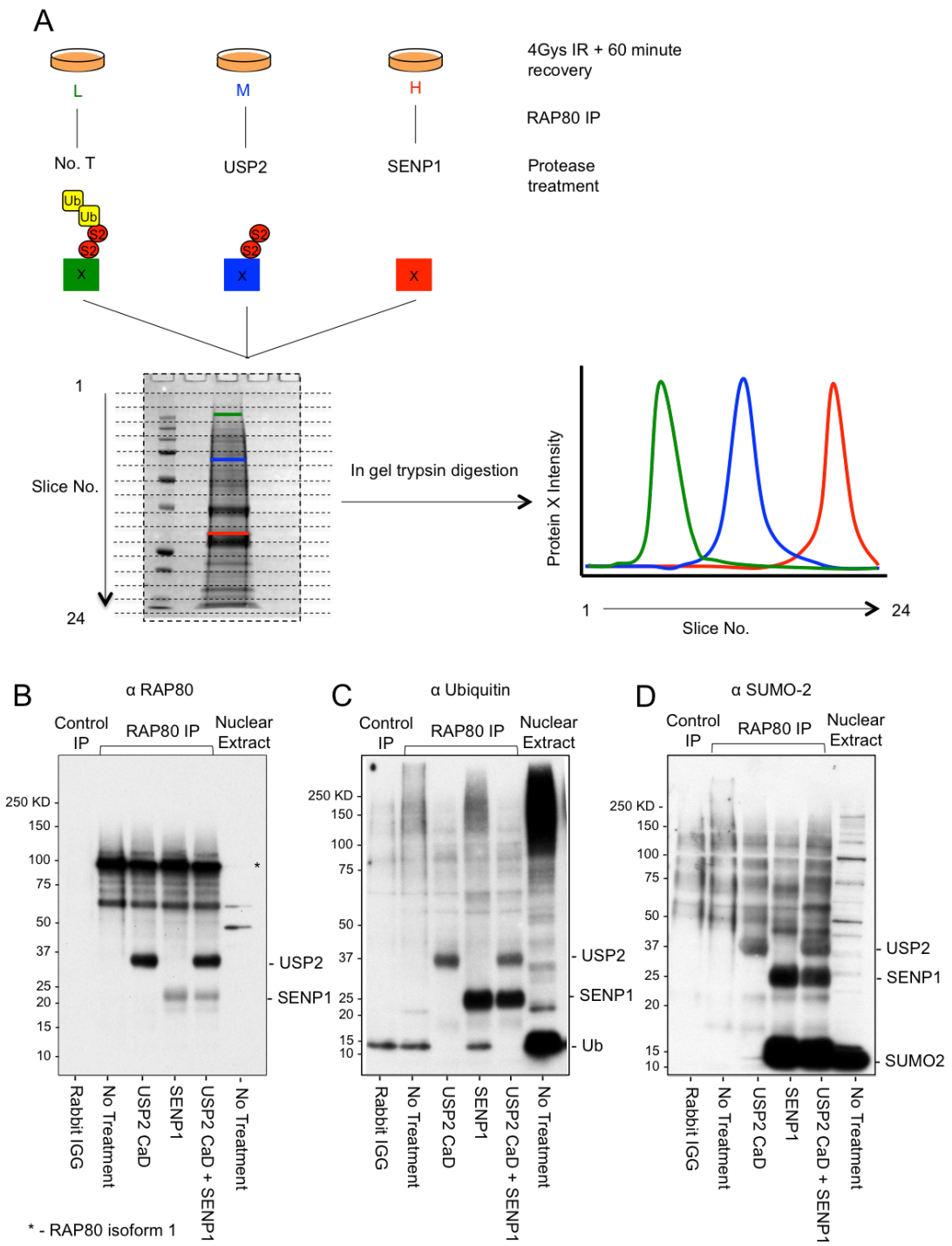


Figure 4.2-12 A SILAC based RAP80 IP gel shift assay utilizing USP2 and SENP1 to evaluate the potential for SUMO-ubiquitin hybrid chain anchoring proteins.

A) Experimental flow diagram. Light, Medium, and Heavy SILAC labeled HEK 293 cells were treated with 4 Gys IR and left to recover for 60 minutes prior to nuclear extract preparation. RAP80 IPs were then performed separately on the Light, Medium, and Heavy nuclear extracts. Protease treatments were then carried on the RAP80 Immunoprecipitated material; the Light labeled RAP80 IP material was not treated with any protease, the Medium labeled

RAP80 material was treated with USP2 CaD, the Heavy labeled material was treated with SENP1.

Western blot analysis of RAP80 Immunoprecipitated material with USP2 CaD, SENP1 and USP2 CaD + SENP1 treatments.

B) RAP80 western blot. RAP80 is efficiently immunoprecipitated from cell extracts under these conditions.

C) Ubiquitin western blot. High molecular weight ubiquitin conjugates are reduced after USP2 CaD treatment, when compared with no treatment control RAP80 IP. SENP1 treatment moves the high molecular weight ubiquitin conjugates lower down the gel, when compared to the no treatment control RAP80 IP. As with the USP2 CaD only treatment, the USP2 CaD + SENP1 treatment reduces the high molecular weight ubiquitin conjugates.

D) SUMO2 western blot. All three protease treatments reduce the high molecular weight SUMO2 conjugates observed in the no treatment control RAP80 IP.

The experimental procedure detailed above was trialled on unlabelled HEK 293 cells, an additional USP2 CaD+SENP1 treatment was also tested along with the single USP2 CaD or SENP1 treatments before western blot analysis was carried out against RAP80, Ubiquitin, and SUMO2 (**Figure 4.2-12, B/C/D**). RAP80 was effectively immunoprecipitated and then detected by western blot in all the RAP80 IP trial conditions (**Figure 4.2-12, B**). As expected high molecular weight ubiquitin conjugates were associated with the No treatment RAP80 IP, were severely reduced in the USP2 CaD treatment RAP80 IP. The high molecular weight ubiquitin conjugates were also observed to shift down the gel in the SENP1 treatment RAP80 IP. As expected the high molecular weight ubiquitin conjugates were again reduced in the USP2 CaD+SENP1 treatment RAP80 IP (**Figure 4.2-12, C**). High molecular weight SUMO conjugates were associated with the No treatment RAP80 IP. Interestingly, the USP2 CaD treatment reduces the observed high molecular weight SUMO modifications observed to be associated with the No treatment RAP80 IP. As expected the SENP1 treatment and the USP2 CaD+SENP1 severely reduced the high molecular weight SUMO modifications associated with the No treatment RAP80 IP (**Figure 4.2-12, D**). Thus both SUMO and ubiquitin modifications of RAP80 associated material are observed to be effected indirectly by the action of the other modifications'

specific protease, USP2 CaD treatment was observed to reduce the apparent size of the high molecular weight SUMO conjugates associated with RAP80, and, SENP1 treatment observed to reduce the size of the high molecular weight ubiquitin modifications associated with RAP80.

HEK 293 cells were grown in Light, Medium, or Heavy SILAC medium, extracts were then prepared and treated in accordance with **2.2.8.4** before the sample were combined and separated by SDS-PAGE. The separated sample was then cut into 24 slices (**Figure 4.2-13, A**). The distance from the top of the gel to the leading edge of the slice was measured and an approximate molecular weight range was given to each slice. This would allow for any proteins observed in this slice to be given an approximate molecular weight and the minimum and maximum differences in molecular weights between slices to be determined. Analysis of all proteins identified in the RAP80 IPs by apparent molecular weight was then conducted comparing the apparent molecular weight of the protein based on its average position (slice) in the gel against the predicted molecular weight of that protein. This analysis suggests that the vast majority of proteins observed in the gel, fall within a +/- 20 kDa range of the predicted molecular weight of that protein (**Figure 4.2-13, B**). This suggests that the majority of proteins observed could not be heavily modified by SUMO and ubiquitin. Apparent molecular weight can also be compared between SILAC labels. Comparison of the apparent molecular weight of proteins from each sample gives the molecular weight difference after USP2 CaD treatment (Medium-Light), and SENP1 treatment (Heavy-Light). The difference in apparent molecular weight for each protein observed after USP2 CaD treatment was then compared to the difference in apparent molecular weight for each protein after SENP1 treatment (**Figure 4.2-13, C**). The distribution of the data suggests that as expected the apparent molecular weight of the vast majority of proteins observed is not altered by more then +/- 20 kDa this suggests that these proteins could not be significantly modified by SUMO or ubiquitin. Proteins that show a decrease in apparent molecular weight after both USP2 CaD and SENP1 treatments are circled in red (**Figure 4.2-13, C).** Surprisingly a large number of proteins are observed to increase in molecular weight after protease treatment. This is likely a technical

discrepancy due to a complication with resolubilsing some of the proteins during the TCA precipitation used to concentrate the samples after protease treatment so that more of each sample could be analysed by SDS-PAGE.

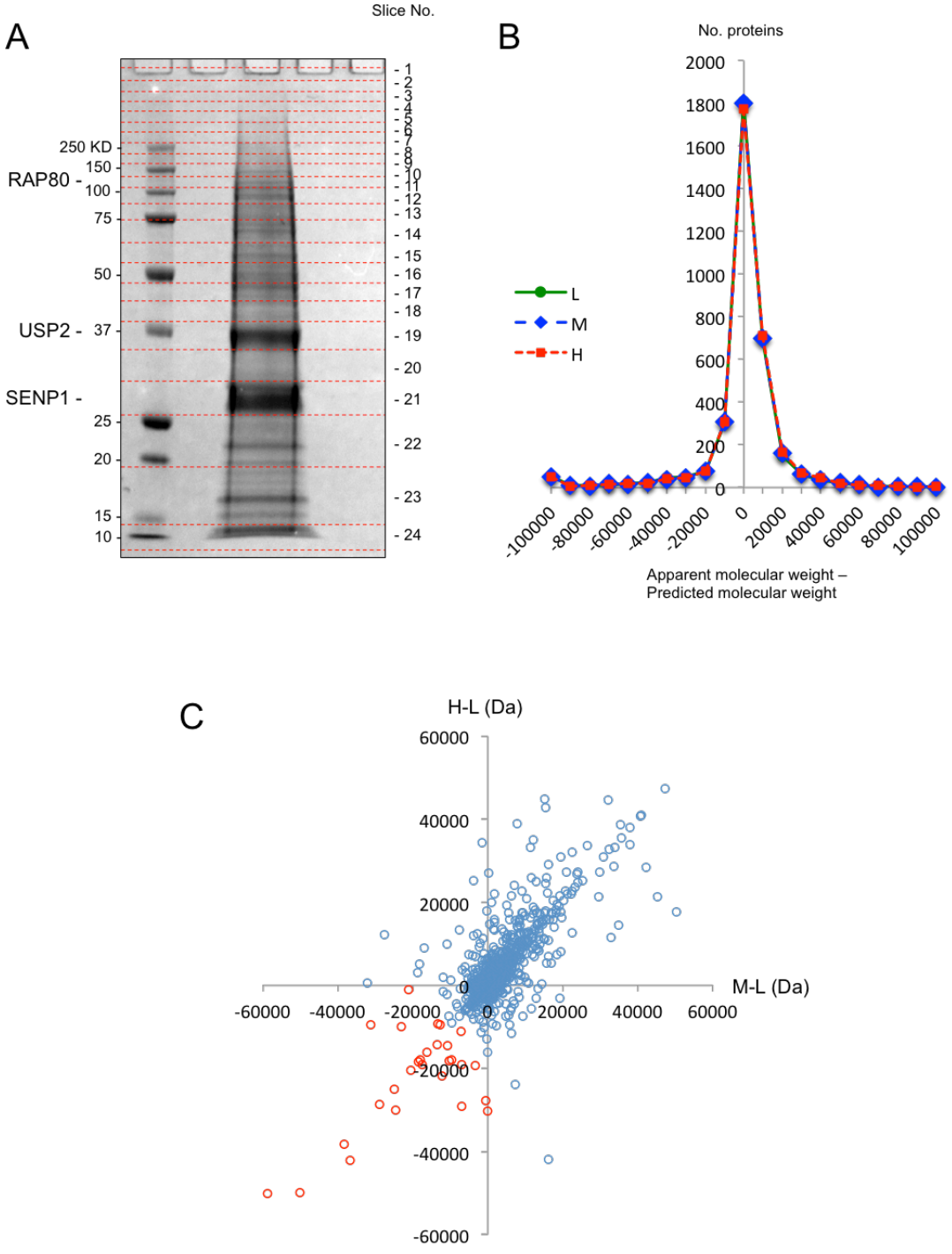


Figure 4.2-13 Analyzing a SILAC based RAP80 IP gel shift assay utilizing USP2 and SENP1 to evaluate the potential for SUMO-ubiquitin hybrid chain anchoring proteins.

A) Coomassie stained gel for Light, Medium, and Heavy labeled RAP80 IP samples. Red dashed lines outline the slices and are numbered appropriately.

B) Molecular weight analysis of proteins found in gel. The majority of proteins are found within 20 kDa of their endogenous molecular weight.
C) Comparison of difference in molecular weight of any given protein by SILAC labeling. The difference in molecular weight is given in daltons. Data in red are proteins that's molecular weight decrease in the USP2 CaD (Medium) treated sample and the SENP1 (Heavy) treated samples when compared to the no treated samples (Light).

The distribution of a protein throughout the gel can then be compared with respect to the three SILAC labelled variants of that protein. This can then be compared to the expected molecular weight of a protein and thus the expected position of that protein in the gel. Light SILAC labelled ubiquitin that had not been treated with a protease is distributed evenly throughout the gel slices (**Figure 4.2-14, A**). This would suggest that as expected, the Light labelled ubiquitin is predominantly conjugated to proteins. The levels of the Medium labelled ubiquitin can be seen to spike in slice 22 and 24, roughly corresponding to the molecular weight of di- or mono-ubiquitin. This suggests that the USP2 CaD protease treatment has resulted the deubiquitylation of proteins and the generation of free ubiquitin. Interestingly, the levels of the Heavy labelled ubiquitin, that had been treated with SENP1, were observed to spike in slice 22, this observation could suggest SENP1 treatment results in the presence of mono-ubiquitylated SUMO species. The levels of Light and Medium labelled SUMO2 are distributed evenly throughout the length of the gel (**Figure 4.2-14, C**). This would suggest that the majority of Light and Medium labelled SUMO2 is predominantly conjugated to proteins. The intensity of Heavy labelled SUMO2 is seen to spike in slice 23 corresponding to the endogenous molecular weight of SUMO2. As expected, treating the Heavy labelled RAP80 IP with SENP1 deconjugated SUMO2 from proteins. A similar spike in slice 23 is seen for the Heavy labelled SUMO1 (**Figure 4.2-14, B**). This suggests that SUMO1 was also deconjugated from proteins with SENP1 treatment.

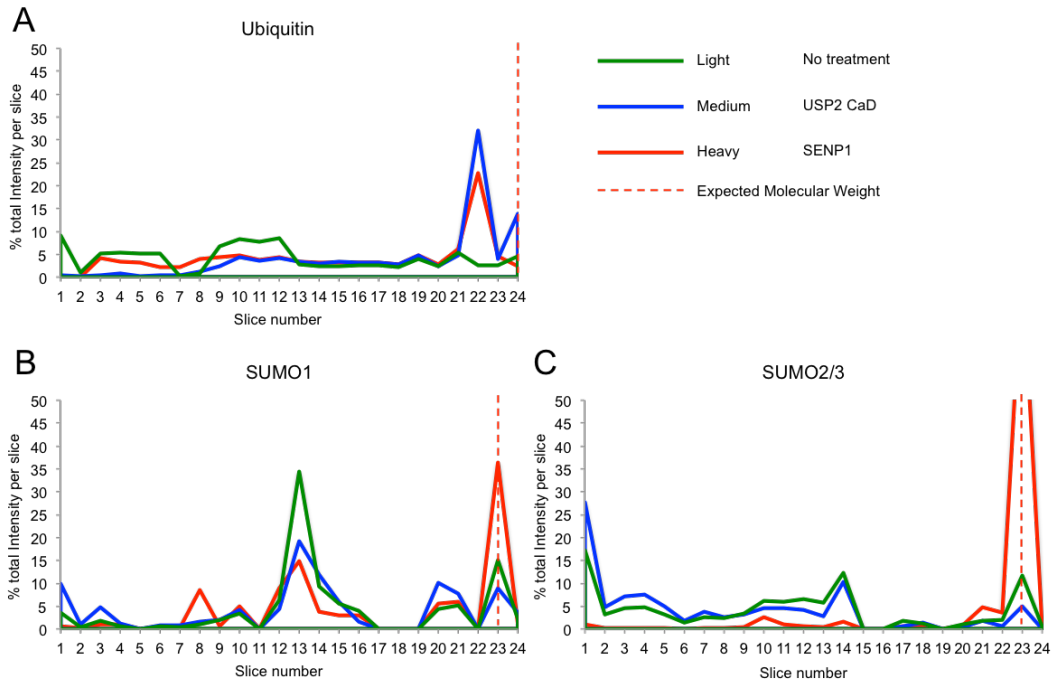


Figure 4.2-14 Evaluating ubiquitin and SUMO interplay via SILAC based RAP80 IP gel shift assay utilizing USP2 and SENP1.

Graphical representation of the percentage of total protein intensity for ubiquitin, SUMO2, and SUMO1 by slice for each SILAC label. Light (no treatment) sample is represented by purple line. Medium (USP2 CaD treatment) sample is represented by the blue line. Heavy (SENP1 treatment) is represented by the orange line. Red dashed line represents the endogenous molecular weight of the protein.

A) Ubiquitin from the Light sample is distributed across the gel, suggesting most ubiquitin is conjugated to proteins. Ubiquitin from the medium sample spikes slice 22 and 24 representing di- and monomeric free ubiquitin. Ubiquitin from the Heavy sample spikes at slice 22.

B) SUMO1 from the light sample spikes at around slice 13 suggesting the majority of SUMO1 is conjugated to proteins. SUMO1 from the medium sample spikes at 13 and 20/21 suggesting that SUMO1 is conjugated to proteins. SUMO1 from the Heavy sample spikes at slice 23 suggesting mostly free SUMO1.

C) SUMO2 from the Light sample is distributed across the gel suggesting conjugated SUMO2. SUMO2 from the Medium sample is distributed across the gel suggesting conjugated SUMO2. SUMO2 from the Heavy sample spikes at slice 23 suggesting mostly free SUMO2.

RAP80 was positively identified in the sample from all three SILAC labeled RAP80 IPs. A RAP80 intensity spike was identified in Slice 12 corresponding well to the predicted molecular weight of isoform 1 of RAP80

(**Figure 4.2-15 top panel**). Three other RAP80 intensity spikes were identified corresponding well with the predicted molecular weight of RAP80 isoforms 2,4 and 5 in slices 15, 18, and 21 respectively. Four other members of the BRCA 1A complex were identified in all SILAC labeled RAP80 IPs. Interestingly, BRCC36, Merit40, Abraxas, and BRE are all identified in slices that correspond well to the predicted molecular weight of the each of the protein. This suggests that these members of the BRCA 1A complex are not heavily modified. Thus, members of the BRCA 1A complex do not appear to be modified with SUMO, ubiquitin, or SUMO-ubiquitin hybrid chains when associated with RAP80.

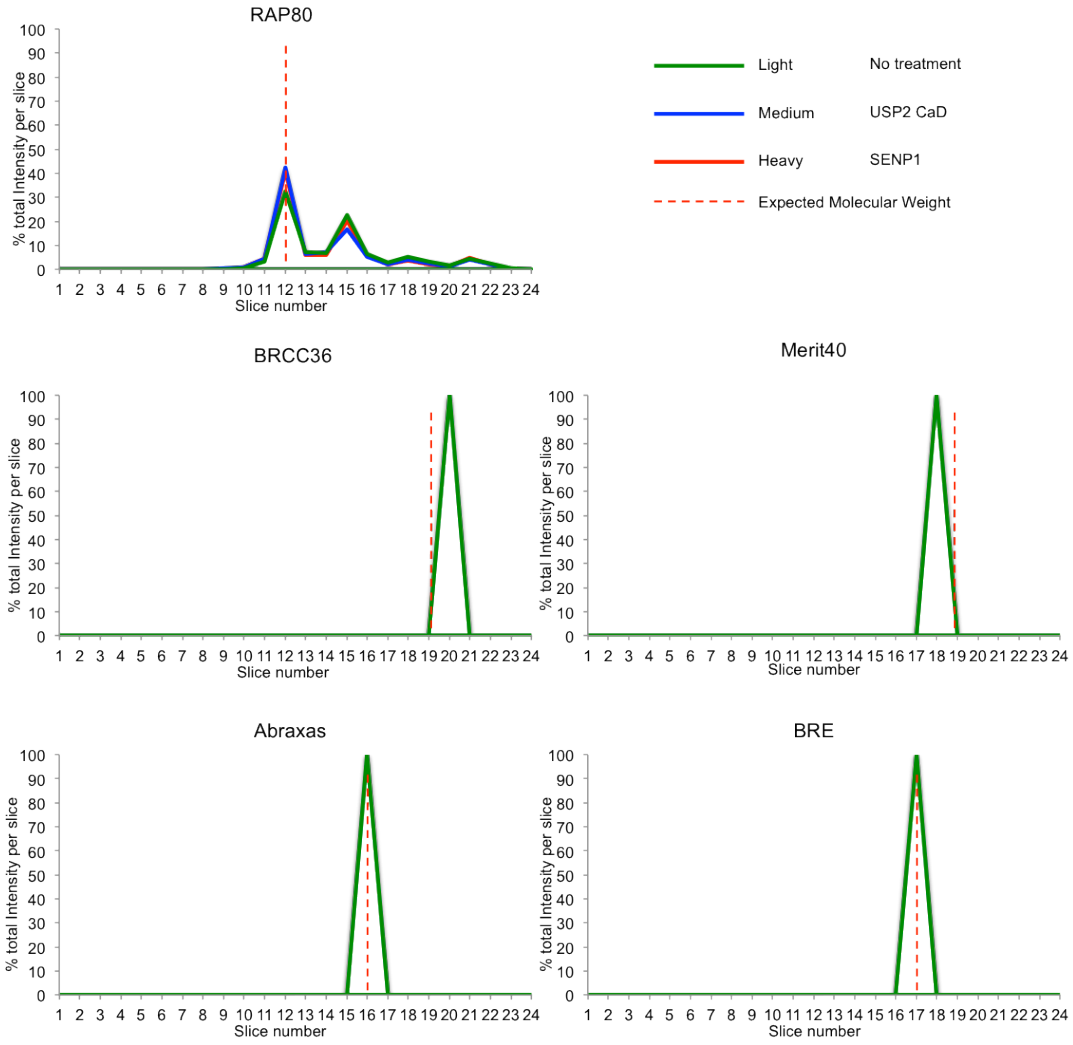


Figure 4.2-15 Evaluating the BRCA 1A complex components via SILAC based RAP80 IP gel shift assay utilizing USP2 and SENP1.

RAP80 spikes at slice 12 for all SILAC labels representing the endogenous weight of RAP80 isoform 1, three additional spikes at 15,18,21 correspond with the predicted molecular weights of the RAP80 isoforms 2, 4, and 5. BRCC36 was identified exclusively in slice 20 for all SILAC labels approximately representing the endogenous weight of BRCC36. Merit40 was identified exclusively in slice 18 for all SILAC labels approximately representing the molecular weight of Merit40. Abraxas was identified exclusively in slice 16 for all SILAC labels representing the endogenous molecular weight of the protein. BRE was identified exclusively in slice 17 representing the endogenous molecular weight of the protein.

---

44 proteins were observed to have lower apparent molecular weight after both USP2 CaD (Medium SILAC label), and SENP1 (Heavy SILAC label) treatments when compared to no treatment (Light SILAC label). Of these, 22 proteins shifted in a pattern consistent with the expected shift generated by the removal of a SUMO-ubiquitin hybrid chain by the USP2 CaD, or SENP1 treatments specified above and the predicted unmodified molecular weight of protein in question (**Figure 4.2-16, A**).

A

Protein Identifier
DSG1
CPS1
LTF
DCD
LYZ
S100A7
S100A8
S100A9
GTPBP5
TG2
JUP
GAPVD1
MED15
KIDINS220
KDM2A
MBNL1
DNMT3A
ELL
PPHLN1
LAGE3
HSP90AB2P

B

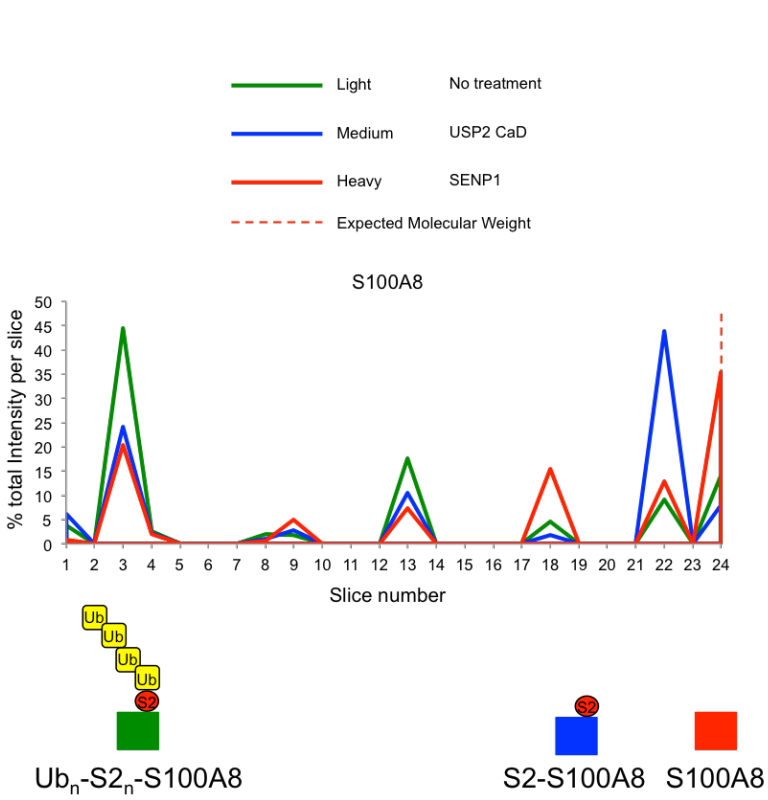


Figure 4.2-16 Potential SUMO-ubiquitin hybrid chain anchoring proteins.

A) List of potential SUMO-ubiquitin hybrid chains anchoring proteins from SILAC based RAP80 immunoprecipitation. Proteins listed have intensity shift pattern associated with sequential removal of ubiquitin and SUMO modification after USP2 CaD, or SENP1 treatment.
B) S100A8 shows intensity pattern of a protein hypermodified by SUMO and ubiquitin. The peak intensity from the light sample is found in slice 3, suggesting a heavily modified version of S100A8. The peak intensity from the USP2 CaD treated Medium sample is found in slice 22, suggesting a modified version of S100A8. The peak intensity of the SENP1 treated Heavy sample is found in slice 24, corresponding to the unmodified molecular weight of S100A8. With a predicted molecular weight of 10KDa, unmodified S100A8 should appear in slice 24 (Red dashed line). Cartoon depicts potential modified species of S100A8 that could explain relative intensity peaks. S100A8 is coloured to match primary intensity peak for each SILAC label. SUMO is represented by red circle. Ubiquitin is represented by yellow boxes.

S100A8, a calcium- and zinc-binding protein, was observed to shift in a pattern consistent with that expected of a protein modified with a hybrid chain of SUMO and ubiquitin (**Figure 4.2-16, B**). The peak intensity of the Light labelled no treated S100A8 is observed in slice 3. This corresponds to a hyper-modified S100A8, depicted below in cartoon form. The peak intensity of the Medium labelled USP2 CaD treated S100A8 is observed in slice 22. This would suggest an apparent molecular weight of between 20-25 KDa, consistent with the predicted molecular weight of a mono SUMOylated S100A8. The peak intensity of the Heavy labelled SENP1 treated S100A8 is observed in slice 24. This corresponds well with the predicted molecular weight of 10 KDa for unmodified S100A8. Thus, S100A8 shifts through the gel in a pattern consistent with that expected of a protein modified by a SUMO-ubiquitin hybrid chain. Unfortunately, it is unknown if S100A8 is SUMO modified, however it does contain a number of lysine residues capable of accepting SUMO modifications, one of which is contained within putative inverted SUMO modification consensus motifs, D32-L33-K34-K35. Therefore further work would be required to establish if S100A8 is indeed modified by SUMO-ubiquitin hybrid chains.

4.3 Discussion

Signalling via hybrid chains of SUMO and ubiquitin would allow for a highly specific response to specific stimuli *in vivo*; due in part to the complex orchestration of numerous enzymes required to create such chains (**Figure 3.1-2**). Recruitment of RAP80 to the sites of DNA damage is conducted, at least in some respect, by the presence of K63 linked poly-ubiquitin chains modifying yH2Ax. However, more recent evidence has suggested that RAP80 is capable of interacting with RNF4 generated SUMO-ubiquitin hybrid chains facilitating the recruitment of RAP80 and the BRCA 1A complex to the sites of DNA repair (Guzzo et al., 2012).

4.3.1 RAP80 has affinity for both SUMO and ubiquitin *in vitro* and *in vivo.*

Guzzo and colleagues suggested that RAP80 has the potential to interact with hybrid chains of SUMO and ubiquitin (Guzzo et al., 2012). The N-terminal region of RAP80 containing a domain composed of a SIM and two UIMs is thought to be capable of recruiting RAP80 to the sites of DNA damage (Guzzo et al., 2012). Here, a truncation of RAP80 containing the tandem SIM UIMUIM domain (RAP80 1-150) was found in the elutate from both SUMO and ubiquitin containing resins types (**Figure 4.2-1, A**). Interestingly, the N-terminal of RAP80 was present in relatively higher levels in the eluate from the poly-SUMO poly-ubiquitin hybrid chain containing resin when compared to the either the mono-ubiquitin poly-SUMO hybrid chain, ubiquitin or SUMO chain containing resin types (**Figure 4.2-1, A**). This corresponds well with the data generated from SUMO-ubiquitin hybrid chain pull down assay using cell extracts, where RAP80 was observed only in the eluate for the hybrid chains consisting of poly-ubiquitin poly-SUMO (**Figure 3.2-8**). Suggesting the poly-ubiquitin poly-SUMO hybrid chains are the preferential interaction substrate for RAP80 in these assays.

It was also reported that RAP80 was recruited to sites of DNA damage due to the activity of RNF4 generating hybrid chains of SUMO and ubiquitin (Guzzo et al., 2012). To determine whether RAP80 did indeed interact with SUMO and ubiquitin in tandem at the sites of DNA damage a *in situ* approach

was taken. The proximity ligation assay is a strong and robust technique developed to allow observations of closely associated pairs of proteins to be determined (Gullberg et al., 2002). Used in conjunction with GFP-tagged RAP80 expressing U2OS cells, PLAs were developed to study the association between RAP80 with SUMO, and RAP80 with ubiquitin (**Figure 4.2-3**). This allowed observations to be made of a close association between RAP80 and SUMO, and RAP80 and ubiquitin *in situ* (**Figure 4.2-3**). As RAP80 was seen to closely associate with both SUMO and ubiquitin, a PLA assay was then developed to determine whether RAP80 is associated with SUMO and ubiquitin in the same foci. Positive signals from the SUMO/ubiquitin PLA were then shown to co-localize with GFP-RAP80 positive foci (**Figure 4.2-4**). Thus, RAP80 appeared to have the potential to associate with both SUMO and ubiquitin at the same time in the same foci. This however does not indicate the presence of SUMO-ubiquitin hybrid chains interacting with RAP80 but suggests that RAP80, SUMO and ubiquitin are all closely associated and gives credence to the hypothesis that RAP80 could be interacting with both SUMO and ubiquitin potentially in the form of a SUMO-ubiquitin hybrid chain.

4.3.2 Is there a role for STUbLs in the recruitment of RAP80 to the sites of DNA damage

If RAP80 is interacting with hybrid chains of SUMO and ubiquitin at the sites of DNA damage, STUbLs such as RNF4 and RNF111 are thought to be the primary candidates capable of building such chains. As such, RAP80 recruitment to the site of DNA damage was explored in CRISPR/CAS9 ΔRNF4 HCT116 cell lines. HCT116 cells are human colorectal cancer cells commonly used in the study of homologous recombination (Huang et al., 2008; Fattah et al., 2010; Mauro et al., 2012). Recruitment of RAP80 appeared to be unaffected in the ΔRNF4 HCT116 cell lines when compared to the parental wild type cell line (**Figure 4.2-6**). After DNA damage, RAP80 is still recruited into yH2Ax positive foci. In the RNF4 postitive parental HCT116 cell line and both ΔRNF4 HCT116 cell line, approximately 45% of cells positive for yH2Ax and RAP80 had 1 or more RAP80 positive yH2Ax foci (**Figure 4.2-6, B**). The average number of positive foci per cell was also unaltered in the ΔRNF4 HCT116 cells when compared to the RNF4 positive

wild type HCT116 cells (Data not shown). Thus in the cell lines used, the complete ablation of RNF4 does not appear to have any affect on the recruitment of RAP80 to the sites of DNA damage. RNF111, a newly identified STUbLs, has also been suggested to have a role in the DDR (Poulsen et al., 2013). Partial knockdown of RNF111 by siRNA had a limited affect in the recruitment of RAP80 to yH2Ax positive foci in either RNF4 positive HCT116 cells or ∆RNF4 HCT116 cell lines (**Figure 4.2-7**). This is an interesting result as it demonstrates that neither RNF4 nor RNF111, the only well described mammalian STUbLs, play a critical role in the recruitment of RAP80 to the sites of DNA repair. This extensively suggests two things; either another, as of yet unidentified STUbL, is important for the recruitment of RAP80 to the sites of DNA repair or perhaps STUbL-generated SUMO-ubiquitin hybrid chains are not critical for the recruitment of RAP80 to the sites of DNA repair.

∆RNF4 HCT116 cell lines used in these experiments contained large amounts of endogenous yH2Ax positive foci under normal growth conditions when compared to the wild type HCT116 cells (**Figure 4.2-8**). These cells however appeared to have no difficulties growing and doubling as cell cycle analysis showed no discrepancy when compared to the wild type cells (**Figure 4.2-8, C**). This suggests anomalies in these cells response to the DDR that may be affecting the results seen. As a result, further analysis of these cells would be recommended. The recruitment of RAP80 to sites of DNA damage could also be analysed after the complete knockout of RNF4 in another cell type commonly used in the study of the DDR such as human U2OS or HEK293 cells.

4.3.3 Are any proteins associated with RAP80 via SUMO ubiquitin hybrid chains

To determine whether the SUMO and ubiquitin observed to be in close association with RAP80 (**Figure 4.2-1-4.2-4**) is important for the recruitment of RAP80 to sites of DNA repair or for the recruitment of other RAP80 associated proteins including other members of the BRCA 1A complex, two

RAP80 IP based experiments were devised. Firstly, RAP80 was immunoprecipitated from HEK293 cells extracts under endogenous and IR treated conditions, then a SENP1 treatment designed to remove endogenous SUMO modifications was applied in the experiment. This experimental set up allowed for proteins interacting with RAP80 in a SUMO specific manner to be identified. Interestingly, western blot analysis of the RAP80 IPs from these extracts, suggested high molecular weight ubiquitin modification were dramatically decreased in the RAP80 associated material from the cell extracts treated with SENP1. This suggests the removal of ubiquitin observed after the SENP1 treatment was associated with RAP80 in a SUMO dependent manner, that the ubiquitin is modifying proteins that also carry SUMO modifications, and that the SUMO modification is required for the interaction of the ubiquitin/SUMO/protein complex to RAP80. Thus, in the context of the association with RAP80, SUMO and ubiquitin are observed to behave in a highly cooperative way. It is interesting to note that the three proteins seen to change most dramatically between the RAP80 IPs from extracts not treated with SENP1, and the RAP80 IPs from extracts treated with SENP1; C5orf25, TOPORs, and RAD54L2 are all known to be SUMO modified. The known SUMO modification of these proteins allows for the potential that these proteins could themselves be modified with hybrid chains of SUMO and ubiquitin. However, some care must be taken with such assumptions as C5orf25, TOPORs, and RAD54L2 are also known to interact with SUMO non-covalently via SUMO interaction motifs. Therefore the SENP1 treatment could be reducing the association of these proteins with RAP80 by reducing the non-covalent interaction directed via SIM interactions of these proteins with the SUMO modifications associated with the RAP80 associated material. The identification of C5orf25 as a SUMO dependent RAP80 associated protein is intriguing. C5orf25, like RAP80 was identified as a potential SUMO-ubiquitin hybrid chain interacting protein in the proteomic data generated by the hybrid chain pull-down assay (**Figure 3.2-7**). If a SUMO-ubiquitin hybrid chain does not directly modify C5orf25, then this could suggest that C5orf25 displays affinity for similar SUMO/ubiquitin modification patterns as RAP80. This would however require further investigation to determine the combinations of SUMO and ubiquitin moieties that show the greatest affinity

for each protein. However, this would not be too surprising given that they both show affinity for the same poly-SUMO poly-ubiquitin hybrid chain type used in the hybrid chain resin pull-down experiment (**Figure 3.2-7**). If RAP80 is indeed precipitating SUMO-ubiquitin hybrid chains it is possible RAP80 is not the only hybrid chain interacting protein associated with such chains.

The second RAP80 immunoprecipitation based experiment was devised using Heavy, Medium, and Light SILAC labels to identify more robustly whether any proteins could be identified as anchoring SUMO-ubiquitin hybrid chains via molecular weight gel shift assays. Firstly, an immunoprecipitation with a RAP80 antibody was used to precipitate RAP80 along with any potential SUMO-ubiquitin hybrid chain bound proteins from cell extracts created from cells grown separately in either Light, Medium, or Heavy labelled growth medium. Subsequently, the Light, Medium, and Heavy labelled RAP80 Immunoprecipitated material were treated with; no protease (no treatment), USP2 CaD, or SENP1 treatments respectively. These treatments were designed to hypothetically leave the hybrid chains unaltered, remove the ubiquitin element of the hybrid chain or remove the entire SUMO-ubiquitin hybrid chain (**Figure 4.2-12, A**). After the samples were combined and analysed by SDS-PAGE, the shift pattern of each protein with respect to SILAC label could be analysed (**Figure 4.2-13**). Analysis of the BRCA 1A complex components for modification by hybrid chains of SUMO and ubiquitin shows no evidence for modification and thus the hypothesis that the BRCA 1A complex could be held together by SUMO-ubiquitin modifications proved unfounded (**Figure 4.2-15**). By this method 22 proteins were identified as shifting through a gel in a pattern consistent with that expected of a protein that was modified by both SUMO and ubiquitin. S100A8 shifted with the expected pattern of a protein modified by a SUMO-ubiquitin hybrid chain (**Figure 4.2-16**). Unfortunately, no data is available that identifies S100A8 to be modified by either SUMO or ubiquitin and to verify this further study would be required. C5orf25, and TOPORs identified in the first RAP80 based IP were not identified in the SILAC based experiment. RAD54L2 was identified in this experiment but was observed in all SILAC labels as an unmodified

protein with no shift through the gel, suggesting RAD54L2 under these conditions is not modified with a SUMO-ubiquitin hybrid chain.

4.3.4 Conclusion

The N-terminal SIM UIMUIM tandem domain containing region of RAP80 (RAP80 1-150) appears to interact with both SUMO and ubiquitin, with a relatively higher degree of affinity shown for poly-SUMO poly-ubiquitin hybrid chains *in vitro* (**Figure 4.2-1, A**). In a cellular context, SUMO, ubiquitin, and RAP80 are all closely associated after DNA damage inducing stimuli (**Figure 4.2-4**). It appears that SUMO and ubiquitin play a cooperative role in their association with RAP80. However, the degree of cooperativity is still unclear. Interestingly, complete knock out of the human STUbL RNF4 showed no effect on the recruitment of RAP80 to the sites of DNA damage (**Figure 4.2-6**), suggesting that RNF4 generated SUMO-ubiquitin hybrid chains are not dependent for the recruitment of RAP80 and subsequently the BRCA 1A complex to the sites of DNA damage as had been previously suggested (Guzzo et al., 2012). Likewise, siRNA knockdown of RNF111 showed no discernable effect on the recruitment of RAP80 to the sites of DNA damage (**Figure 4.2-7**). Three known SUMO modified protein; C5orf25, TOPORs, and RAD54L2 were observed to interact with RAP80 associated material in a SUMO dependent manner, leading to the possibility these proteins could be modified by hybrid chains of SUMO and ubiquitin (**Figure 4.2-11**). However, because all three proteins also contain SIMs the data was unable to discriminate between SUMO dependent non-covalent interactions with SUMO modified version of these proteins and RAP80 associated material or between these proteins and SUMO modified RAP80 associated material. Utilising triple SILAC labelling with USP2 CaD and SENP1 treatment of RAP80 IPs allowed the development of an assay to detect SUMO-ubiquitin hybrid chain anchoring proteins (**Figure 4.2-12**). Interesting, 22 proteins were identified from the RAP80 IPs that migrate through the gel consistent with the expected migration pattern of proteins modified with SUMO and ubiquitin (**Figure 4.2-16**). S100A8 was identified as shifting through the gel with a pattern expected of a protein modified by a SUMO-ubiquitin hybrid chain (**Figure 4.2-16**). However, further evidence would be required to discern whether S100A8 is

modified by SUMO-ubiquitin hybrid chains. Also of note, during the SILAC mass spectrometry analysis no SUMO peptides were identified carrying the characteristic GlyGly motif that indicates ubiquitylation. Thus it is difficult to definitively say that SUMO-ubiquitin hybrid chains modified those proteins or are even associated with RAP80. Further optimisation of this experimental procedure would be required to resolve this issue.

5 Conclusions

Since the identification of the ubiquitin conjugation pathway almost 30 years ago, the Ubl field has grown dramatically to become an important area for understanding many essential cellular processes. Recent published data and evidence presented in this thesis has shown that the once thought distinct ubiquitin and SUMO conjugation pathways converge during several vital cellular functions.

In the first part of this thesis, a protocol was designed to identify proteins with the potential to interact in a specific manner with hybrid chains of SUMO and ubiquitin via affinity chromatography coupled with high-resolution mass spectrometry. This was successful, identifying around 30 proteins that interacted specifically with hybrid chains of SUMO and ubiquitin (**Figure 3.2-7**). During the course of these experiments RAP80 was identified to contain an N-terminal region containing a tandem SIM UIMUIM domain that is capable of specifically interacting with K63 linked poly-ubiquitin modified SUMO2 (Guzzo et al., 2012). Gratifyingly, RAP80 was also identified in our attempts to identify SUMO-ubiquitin hybrid chain interacting proteins (**Figure 3.2-7**). Additionally, RAP80 was identified with six other members of the BRCA 1A complex suggesting that this complex may as a whole have affinity for SUMO-ubiquitin hybrid chains (**Figure 3.2-8**). A primary sequence search of proteins known to contain UIMs for SIMs was also carried out (**Figure 3.2-12**). Interestingly, if not unsurprisingly, several proteins were identified as having both UIMs and SIMs by this method were also shown to be potential SUMO-ubiquitin hybrid chain interacting proteins via our affinity chromatography coupled with high-resolution mass spectrometry. This both validated our primary approach and reinforced the concept that proteins critical for many different cellular functions have the domains necessary to recognise both SUMO and ubiquitin and could act as points of convergence between the ubiquitin and SUMO pathways.

In the second part of this thesis the potential interaction of RAP80 with hybrid chains of SUMO and ubiquitin was explored in a cellular context

focused around the sites of DSB after DNA damage. A PLA was developed to observe closely associated SUMO and ubiquitin molecules in RAP80 positive foci, suggesting that a high level of cooperativity exists between SUMO and ubiquitin modifications associated with RAP80 in a cellular context (**Figure 4.2-4**). This was also observed when RAP80 IPs were carried out on cell extracts that had either been treated with SENP1 or left untreated (**Figure 4.2-9**). As expected SENP1 treatment of RAP80 immunoprecipitated material displayed a decrease in high molecular weight SUMO conjugated material when compared to RAP80 imunnoprecipitated material that was left untreated. Interestingly, SENP1 treatment of RAP80 immunopreciptated material resulted in a pronounced decrease in the levels of high molecular weight ubiquitin conjugates, suggesting this material was ubiquitinated in a SUMOylation dependent manner (**Figure 4.2-9**). SENP1 treatment of RAP80 immunoprecipited material also uncovered a handful of proteins that appear to be associated with RAP80 in a SUMO dependent manner, including C5orf25 a proteins that was also identified to show affinity specifically for SUMO-ubiquitin hybrid chains (**Figure 4.2-11**). Importantly the STUbL RNF4 was observed to not be required for the recruitment of RAP80 to the sites of DNA damage, as in ΔRNF4 cell lines, RAP80 was observed at the sites of DNA damage (**Figure 4.2-6**). Finally, the apparent affinity of RAP80 for hybrid chains of SUMO and ubiquitin was utilised in an attempt identify proteins that may anchor hybrid chains of SUMO and ubiquitin (**Figure 4.2-12**). This experimental approach showed promise and successfully identified a handful of proteins with gel shift patterns consistent with that expects of proteins modified with a hybrid chain of SUMO and ubiquitin (**Figure 4.2-16**). This included S100A8, however, further optimization of the experimental technique would be required to confirm that S100A8 was indeed linked to a hybrid chain of SUMO and ubiquitin. Whether hybrid chains of SUMO and ubiquitin are critically important for the recruitment of RAP80 to the sites of DNA damage is still up for debate. RNF4, the primary candidate for generating SUMO-ubiquitin hybrid chains at the sites of damage does not appear to be required for the recruitment of RAP80 to the sites of DNA damage. However, clearly a high degree of cooperativity exists between SUMO, ubiquitin, and RAP80. This cooperativity does not appear to extend to RNF4 generated SUMO-

ubiquitin hybrid chains in the case for RAP80 recruitment to the sites of damage, but this does not rule out that these chains maybe important for other cellular processes.

The possibility of SUMO-ubiquitin hybrid chains acting as a signalling mechanism distinct from homotypic chains of SUMO and ubiquitin is of great interest to the SUMO, ubiquitin and greater PTM fields. Recent advances in mass spectrometry analysis has allowed for the identification of SUMO modified by ubiquitin and vice versa, posing the question as to whether the PTM of proteins that themselves act as PTM could produce hybrid signal of both PTMs? The generation of substrates that are both SUMOylated and ubiquitylated requires the recruitment and activation of enzymes from both SUMO and ubiquitin conjugation pathways. Thus, a hybrid signal generated by the orchestration of both the SUMO and ubiquitin conjugation pathways would inherently be more specific than that of either the SUMO or ubiquitin conjugation pathways alone. Protein specificity for these hybrid signals would also be increased as interaction elements for both SUMO and ubiquitin would be required to work in tandem, as seen in the N-terminal SIM UIMUIM domain of RAP80. The vast number of different combination of SUMO-ubiquitin hybrid chains that could potentially be created, could act as a new reservoir of as yet untapped specific signals used to pinpoint protein complexes to very specific cellular events, bypassing the landscape of more commonly found SUMO or ubiquitin modifications in the cell under normal conditions. In this respect it is also interesting to note that STUbLs have been evolutionary conserved from yeast to humans suggesting some biological pressure to conserve the action of these proteins in cells. Thus, further work to identify potential biological roles for SUMO-ubiquitin hybrid chains would be of great interest to the Ubl field as a whole.

Appendix

Appendix I

2.1.1 General buffer list

Name	Components
6X DNA loading buffer	60 mM Tris-HCL, 50 mM EDTA, 1%(w/v) SDS, 30% (w/v) glycerol, 0.1% (w/v) bromophenolblue, pH 8.0
6X SDS sample buffer	60 mM Tris-HCL, 1%(w/v) SDS, 10%(v/v) glycerol, 100 mM DTT, 0.005 mM bromophenolblue, pH 6.8
Coomassie Blue stain	4mg of Coomassie Blue R25 was dissolved in 1000ml of Ethanol. 200ml of Acetic acid is then added with 800ml of distilled H_2O. Solution is then stirred overnight and then filtered through filter paper.
Destain 1	40% (v/v) EtOH, 7% (v/v) acetic acid
Destain 2	5% (v/v) EtOH, 7% (v/v) acetic acid
MES (1X)	50 mM MES, 50 mM Tris, 0.1% SDS, 1 mM EDTA pH 7.7
MOPs (1x)	50 mM MOPs, 50 mM Tris, 0.1% SDS, 1 mM EDTA pH 7.3
PBS(T)	Phosphate-buffered saline (0.1% Tween)
Reaction buffer	50 mM Tris pH 7.5, 150 mM NaCl, 0.1% NP-40
Tris/Gly transfer buffer	25 mM Tris, 192 mM Glycine 20% EtOH
TBS(T)	Tris-buffered saline (0.1% Tween)
Western blot stripping Buffer	0.2 M NaOH, 0.5 M NaCl

2.1.2 Antibodies list

Antibody	Species	Application	Source
53BP1	Mouse	WB/IF	Bethyl Laboratories
BRCA1	Mouse	WB/IF	Santa Cruz Biotech

Morc3	Rabbit	WB/IP	Santa Cruz Biotech
RAP80	Rabbit	WB/IF/IP/PLA	Bethyl Laboratories
SENP1	Rabbit	WB/IP	Cell Signalling
SUMO2	Rabbit	WB/IF/PLA	Invitrogen
SUMO2	Sheep	WB	Hay Lab stock
Ubiquitin	Rabbit	WB	Invitrogen
Ubiquitin (FK2)	Mouse	WB/IF/PLA	BIOMOL
USP2	Rabbit	WB/IP	Cell signalling
USP28 (m)	Rabbit	WB	Abcam
USP28 (p)	Rabbit	IP	Bethyl Laboratories
yH2A.x	Mouse	WB/IF	MerekMillipore

2.1.3 PCR primer list

Morc 3 Full Length & Mutants

- Full length F - TTTTTTGGATCCATGGCGGCGCAGCCACCCCGCG
- Full length R - TTTTTTCTCGAGTTAAGTACTACTGATTTCACTCATTTGTTCACT
- Morc3 A F - TTTTTTGGATCCATGGCGGCGCAGCCACCCCGCG
- Morc3 A R - TTTTTTCTCGAGTTATTCAACTGGCAAATTTAGAGGATATTCTGT
- Morc3 B F - TTTTTTGGATCCATGGCGGCGCAGCCACCCCGCG
- Morc3 B R - TTTTTTCTCGAGTTAGGGATGTACCAAATCCTCATCTTCAGGTTC
- Morc3 C F - TTTTTTGGATCCGAAGATATACAGAAGCGTCCTGATCAGACATGG
- Morc3 C R - TTTTTTCTCGAGTTAAGTACTACTGATTTCACTCATTTGTTCACT
- Morc3 D F - TTTTTTGGATCCCCCACTTATGAAAAAACCTACAAAAAGACCAAC

- Morc3 D R - TTTTTTCTCGAGTTAAGTACTACTGATTTCACTCATTTGTTCACT
- Morc3 SIM Mutant 1 F - ATAAAAGCGGAGCATGCAGCAGCAGCAGTGGCATTCAACAAGC
- Morc3 SIM Mutant 1 R - GCTTGTTGAATGCCACTGCTGCTGCTGCATGCTCCGCTTTTAT
- Morc3 SIM Mutant 2 F - GATGATGATGGAGATGCAGCAGCAGCAGAAGAAAACAGTACCCC
- Morc3 SIM Mutant 2 R - GGTACTGTTTTCTTCTGCTGCTGCTGCATCTCCATCATCATCATCACC
- Morc3 SIM Mutant 3 F - GAACTGAGAAACCAGCTACTCCTTGTCACTGAGGAAAAAGAG
- Morc3 SIM Mutant 3 R - CTCTTTTTCCTCAGTTGCTGCTGCTGCCTGGTTTCTCAGTTC
- Morc3 SIM Mutant 4 F - TCAGCAAGTGAATTACGATGCAGCAGCAGCAGATGAGATTTTAGGACAAG
- Morc3 SIM Mutant 4 R - CTTGTCCTAAAATCTCATCTGCTGCTGCTGCATCGTAATTCACTTGCTGA

USP2 Catalytic Domain

- CaD F - TTTTTTAAGCTTGCAATTCTAAGAGTGCCCAGGGTC
- CaD R - TTTTTTCTCGAGTTACATTCGGGAGGGCGGG

USP28 Full Length & Mutants

- Full length F - TTTTTTGGATCCATGACTGCGGAGCTGCAGC
- Full length R - TTTTTTAAGCTTTTATTTCACTGTCACAGTTGAAACTCCCTA

- SIM-UIM F - TTTTTTGGATCCATGACTGCGGAGCTGCAGC
- SIM-UIM R -
- TTTTTTAAGCTTTTAATTGGGGTTTTCTCCCCAGACTTCACA
- SIM-ΔUIM F - TTTTTTGGATCCATGACTGCGGAGCTGCAGC
- SIM-ΔUIM R - TTTTTTAAGCTTTTAATGTTTGTTATCATGAGTAAGGTCTAT
- ΔSIM-UIM F - GCCGACGCCACTCATGATAACAAAGATGATCTTCAGGCTGCCATTG
- ΔSIM-UIM R - TTTTGCAATACTTCCTTGTTGGCAGCACTCCCCTCTACTTCAGATGG

2.1.4 Reagent Kit list

Kits	Source
Benzoase	Merekmillipore
Duolink® Using PLA® In SITU Red Starter Kit Mouse/Rabbit	Sigma-Aldrich
DC-Assay	Bio-Rad
ECL Plus	Thermo Scientific
KOD DNA Polymerase	Toyobo Life Sciences
Phusion HF DNA Polymerase	Life Technologies
QIAprep Miniprep Kit	Qiagen
QIAprep PCR Purification	Qiagen
QIAqick Gel Extraction	Qiagen
Restriction Enymes (various)	NEB
RNAiMAX	Life Technologies
T4 DNA Ligase	NEB
Novex NuPAGE 10% Bis-tris gels	Life Technologies
Novex NuPAGE 4-12% Bis-tris gels	Life Technologies

2.1.5 Plasmids & Vectors

Plasmid	Source
pcDNA3.1+HA-C5orf25	DSTT
pET-6His_TEV-RNF168	DSTT
pGEX-6P-USP2	DSTT
pGEX-6P-USP28	DSTT
pHISTEV30a	Hay Lab Stocks
pHISTEV30a -Ub-4xSUMO2ΔN11	Dr. A Plechanovova
pHISTEV30a-4xSUMO2ΔN11	Dr. A Plechanovova
pLou3-MBP-Ube2v2	Dr. A Plechanovova
pSC-HA-Morc3	DSTT

2.1.6 Recombinant Proteins

Protein	Source
4xSUMO2	Thesis author
GST-UIM dimer	DSTT
K63 pUb	DSTT
Morc3	Thesis author
Morc3 A/B/C/D	Thesis author
pUb-4xSUMO2	Thesis author
RNF4	Hay lab stocks
SENP1	Dr. L Shen
Tri K48 Ub	DSTT
Ub-4xSUMO2	Thesis author
Uba1	Hay lab stocks
Ubc13	Hay lab stocks
Ubiquitin	Hay lab stocks
USP2 CaD (Asn259-Met605)	Thesis author
USP28	Thesis author
USP28 N-terminal Mutants	Thesis author

Appendix II

Full list of all proteins characterised with a given specificity for polymeric chains of SUMO, ubiquitin, or hybrid chains of SUMO and ubiquitin from the hybrid chain pull down assay after mass spectrometry analysis.

Gene Name	Predicted Specificity
TULP4	Ub poly-SUMO
ZMYM2	Ub poly-SUMO
ASCC3	poly-Ub poly-SUMO
FAM175A	poly-Ub poly-SUMO
BRE	poly-Ub poly-SUMO
UIMC1	poly-Ub poly-SUMO
BRCA1	poly-Ub poly-SUMO
BABAM1	poly-Ub poly-SUMO
CENPE	poly-Ub poly-SUMO
CCDC50	poly-Ub poly-SUMO
HUWE1	poly-Ub poly-SUMO
RNF168	poly-Ub poly-SUMO
EPS15L1	poly-Ub poly-SUMO
EPN1	poly-Ub poly-SUMO
EPN2	poly-Ub poly-SUMO
BRCC3	poly-Ub poly-SUMO
MORC3	poly-Ub poly-SUMO
MORC4	poly-Ub poly-SUMO
SQSTM1	poly-Ub poly-SUMO
TUBB6	poly-Ub poly-SUMO
ZNF451	poly-Ub poly-SUMO
AHCTF1	Any Ub poly-SUMO
CPLX3	Any Ub poly-SUMO
INTS7	Any Ub poly-SUMO
OBSCN	Any Ub poly-SUMO
C5orf25	Any Ub poly-SUMO

Gene Name	Predicted Specificity
PPIG	Any poly-SUMO
PIN1	Any poly-SUMO
POLDIP3	Any poly-SUMO
PRPF8	Any poly-SUMO
ISY1	Any poly-SUMO
BCAS2	Any poly-SUMO
PLOD1	Any poly-SUMO
PFN1	Any poly-SUMO
P$HB	Any poly-SUMO
PDIA6	Any poly-SUMO
MAGOHB	Any poly-SUMO
PML	Any poly-SUMO
SEC13	Any poly-SUMO
NSUN4	Any poly-SUMO
RAVER1	Any poly-SUMO
RBM8A	Any poly-SUMO
MSI1	Any poly-SUMO
SENP1	Any poly-SUMO
SENP2	Any poly-SUMO
CDC42BPB	Any poly-SUMO
PPP1R10	Any poly-SUMO
PGAM5	Any poly-SUMO
SRP14	Any poly-SUMO
SNRPF	Any poly-SUMO
SNRPG	Any poly-SUMO
SNRPD3	Any poly-SUMO
SUMO3	Any poly-SUMO
ATP1A1	Any poly-SUMO
SF£B5	Any poly-SUMO
UBE2I	Any poly-SUMO
ALFREF	Any poly-SUMO
THRAP3	Any poly-SUMO
TECR	Any poly-SUMO
TRA2A	Any poly-SUMO
TFIP11	Any poly-SUMO
PTPN1	Any poly-SUMO
SNRBP2	Any poly-SUMO
USP39	Any poly-SUMO
SNRNP200	Any poly-SUMO
SNRNP40	Any poly-SUMO
WDR82	Any poly-SUMO
SMU1	Any poly-SUMO
ZC3H14	Any poly-SUMO
FHL1	Any poly-SUMO

Gene Name	Predicted Specificity
EFTUD2	Any poly-SUMO
EFTUD2	Any poly-SUMO
HIBCH	Any poly-SUMO
RPS24	Any poly-SUMO
RPS7	Any poly-SUMO
RPL14	Any poly-SUMO
RPL26	Any poly-SUMO
RPL30	Any poly-SUMO
RPL34	Any poly-SUMO
MCM3AP	Any poly-SUMO
FAHD1	Any poly-SUMO
CLPX	Any poly-SUMO
DDX42	Any poly-SUMO
BSG	Any poly-SUMO
ARRB2	Any poly-SUMO
KCTD12	Any poly-SUMO
CD2BP2	Any poly-SUMO
CHD1L	Any poly-SUMO
CPSF2	Any poly-SUMO
CTPS	Any poly-SUMO
POLR2A	Any poly-SUMO
DPM1	Any poly-SUMO
DDOST	Any poly-SUMO
DNAH6	Any poly-SUMO
ERP29	Any poly-SUMO
ERH	Any poly-SUMO
EIF1AX	Any poly-SUMO
EIF2S1	Any poly-SUMO
TDG	Any poly-SUMO
HNRPDL	Any poly-SUMO
HIST1H2AB	Any poly-SUMO
IFT74	Any poly-SUMO
LMNB1	Any poly-SUMO
LMNB2	Any poly-SUMO
LANCL1	Any poly-SUMO
PGRMC1	Any poly-SUMO
NDUFAF4	Any poly-SUMO
FDXR	Any poly-SUMO
POM121C	Any poly-SUMO
NUP107	Any poly-SUMO
NUP133	Any poly-SUMO
NUP160	Any poly-SUMO
NUP85	Any poly-SUMO

Gene Name	Predicted Specificity
PDE12	K63 poly-Ub
ACAT2	K63 poly-Ub
ANKRD13A	K63 poly-Ub
ANKRD13D	K63 poly-Ub
DDX54	K63 poly-Ub
DNAJB1	K63 poly-Ub
LACRT	K63 poly-Ub
HNRNPA1	K63 poly-Ub
HNRNPA2B1	K63 poly-Ub
IGHA1	K63 poly-Ub
IGHG1	K63 poly-Ub
IGKC	K63 poly-Ub
IGJ	K63 poly-Ub
IGLL5	K63 poly-Ub
LTF	K63 poly-Ub
LCN1	K63 poly-Ub
LYZ	K63 poly-Ub
NCL	K63 poly-Ub
PIGR	K63 poly-Ub
CTR9	K63 poly-Ub
FUS	K63 poly-Ub
TAF15	K63 poly-Ub
TUBB8	K63 poly-Ub
USP37	K63 poly-Ub
USP5	K63 poly-Ub
RPS27A	K63 poly-Ub
UBA1	K63 poly-Ub
MYO6	K63 poly-Ub
C7orf57	K63 poly-Ub
LMNA	K63 poly-Ub
HDAC6	K63 poly-Ub
PA1G4	K63 poly-Ub
TUBGCP6	K63 poly-Ub
MFAP1	K63 poly-Ub
RCN2	K63 poly-Ub
ACAA2	K63 poly-Ub
GLUD1	K63 poly-Ub

References

Bohm, S., M. J. Mihalevic, et al. (2015). "Disruption of SUMO-targeted ubiquitin ligases Slx5-Slx8/RNF4 alters RecQ-like helicase Sgs1/BLM localization in yeast and human cells." DNA Repair 26: 1-14.

Bailey, D. and P. O'Hare (2004). "Characterization of the Localization and Proteolytic Activity of the SUMO-specific Protease, SENP1." Journal of Biological Chemistry 279(1): 692-703.

Boggio, R., R. Colombo, et al. (2004). "A mechanism for inhibiting the SUMO pathway." Molecular Cell 16(4): 549-561.

Bohren, K. M., V. Nadkarni, et al. (2004). "A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus." Journal of Biological Chemistry 279(26): 27233-27238.

Brady, M., N. Vlatkovi*ƒá*, et al. (2005). "Regulation of p53 and MDM2 activity by MTBP." Molecular and Cellular Biology 25(2): 545-553.

Branigan, E., A. Plechanovov√°, et al. (2015). "Structural basis for the RING-catalyzed synthesis of K63-linked ubiquitin chains." Nature Structural and Molecular Biology 22(8): 597-602.

Bruderer, R., M. H. Tatham, et al. (2011). "Purification and identification of endogenous polySUMO conjugates." EMBO Reports 12(2): 142-148.

Burroughs, A. M., M. Jaffee, et al. (2008). "Anatomy of the E2 ligase fold: Implications for enzymology and evolution of ubiquitin/Ub-like protein conjugation." Journal of Structural Biology 162(2): 205-218.

Cantor, S. B., D. W. Bell, et al. (2001). "BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function." Cell 105(1):

149-160.

Chapman, J. R., M. R. G. Taylor, et al. (2012). "Playing the End Game: DNA Double-Strand Break Repair Pathway Choice." Molecular Cell 47(4): 497-510.

Chung, S. S., B. Y. Ahn, et al. (2010). "Control of adipogenesis by the SUMO-specific protease SENP2." Molecular and Cellular Biology 30(9): 2135-2146.

Ciehanover, A., Y. Hod, et al. (1978). "A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes." Biochemical and Biophysical Research Communications 81(4): 1100-1105.

Cox, J. and M. Mann (2008). "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification." Nature Biotechnology 26(12): 1367-1372.

Demuc, A., A. Bosch-Comas, et al. (2009). "The UBA-UIM domains of the USP25 regulate the enzyme ubiquitination state and modulate substrate recognition." PLoS ONE 4(5).

Desterro, J. M. P., J. Thomson, et al. (1997). "Ubch9 conjugates SUMO but not ubiquitin." FEBS Letters 417(3): 297-300.

Dikic I, Wakatsuki S, Walters KJ. (2009) “Ubiquitin-binding domains—from structures to functions.” Nat. Rev. Mol. Cell Biol. 10:659–71

Doil, C., N. Mailand, et al. (2009). "RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins." Cell 136(3): 435-446.

Duprez, E., A. J. Saurin, et al. (1999). "SUMO-1 modification of the acute promyelocytic leukaemia protein PML: Implications for nuclear localisation." Journal of Cell Science 112(3): 381-393.

Eddins MJ, Varadan R, Fushman D, Pickart CM, Wolberger C. (2007). “Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH.” J. Mol. Biol. 367:204–11

Falck, J., J. Coates, et al. (2005). "Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage." Nature 434(7033): 605-611.

Fattah, F., Lee, E. H., Weisensel, N., Wang, Y., Lichter, N., & Hendrickson, E. A. (2010). Ku Regulates the Non-Homologous End Joining Pathway Choice of DNA Double-Strand Break Repair in Human Somatic Cells. *PLoS Genetics*,*6*(2), e1000855.

Finley, D., B. Bartel, et al. (1989). "The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis." Nature 338(6214): 394-401.

Galanty, Y., R. Belotserkovskaya, et al. (2012). "RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair." Genes and Development 26(11): 1179-1195.

Galanty, Y., R. Belotserkovskaya, et al. (2009). "Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks." Nature 462(7275): 935-939.

Geoffroy, M. C. and R. T. Hay (2009). "An additional role for SUMO in ubiquitin-mediated proteolysis." Nature Reviews Molecular Cell Biology 10(8): 564-568.

Geoffroy, M. C., E. G. Jaffray, et al. (2010). "Arsenic-induced SUMO-dependent recruitment of RNF4 into PML nuclear bodies." Molecular Biology of the Cell 21(23): 4227-4239.

Gill, G. (2003). "Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity." Current Opinion in Genetics and Development 13(2): 108-113.

Girdwood, D. W. H., M. H. Tatham, et al. (2004). "SUMO and transcriptional regulation." Seminars in Cell and Developmental Biology 15(2): 201-210.

Goldstein, G., M. Scheid, et al. (1975). "Isolation of a polypeptide that has lymphocyte differentiating properties and is probably represented universally in living cells." Proceedings of the National Academy of Sciences of the United States of America 72(1): 11-15.

Golebiowski, F., I. Matic, et al. (2009). "System-wide changes to sumo modifications in response to heat shock." Science Signaling 2(72).

Gong, L., S. Millas, et al. (2000). "Differential regulation of sentrinized proteins by a novel sentrin- specific protease." Journal of Biological Chemistry 275(5): 3355-3359.

Gong, L. and E. T. H. Yeh (2006). "Characterization of a family of nucleolar SUMO-specific proteases with preference for SUMO-2 or SUMO-3." Journal of Biological Chemistry 281(23): 15869-15877.

Gudjonsson, T., M. Altmeyer, et al. (2012). "TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes." Cell 150(4): 697-709.

Gullberg, M., S. Fredriksson, et al. (2003). "A sense of closeness: Protein detection by proximity ligation." Current Opinion in Biotechnology 14(1): 82-86.

Guzzo, C. M., C. E. Berndsen, et al. (2012). "RNF4-dependent hybrid SUMO-ubiquitin chains are signals for RAP80 and thereby mediate the recruitment of BRCA1 to sites of DNA damage." Science Signaling 5(253).

Guzzo, C. M. and M. J. Matunis (2013). "Expanding SUMO and ubiquitin-mediated signaling through hybrid SUMO-ubiquitin chains and their receptors." Cell Cycle 12(7): 1015-1017.

Guzzo, C. M., A. Ringel, et al. (2014). "Characterization of the SUMO-binding activity

of the myeloproliferative and mental retardation (MYM)-type zinc fingers in ZNF261 and ZNF198." PLoS ONE 9(8).

Haindl, M., T. Harasim, et al. (2008). "The nucleolar SUMO-specific protease SENP3 reverses SUMO modification of nucleophosmin and is required for rRNA processing." EMBO Reports 9(3): 273-279.

Hands, K. J., D. Cuchet-Lourenco, et al. (2014). "PML isoforms in response to arsenic: High-resolution analysis of PML body structure and degradation." Journal of Cell Science 127(2): 365-375.

Hannoun, Z., S. Greenhough, et al. (2010). "Post-translational modification by SUMO." Toxicology 278(3): 288-293.

Hay, R. T. (2001). "Protein modification by SUMO." Trends in Biochemical Sciences 26(5): 332-333.

Hay, R. T. (2004). "Modifiying NEMO." Nature Cell Biology 6(2): 89-91.

Hay, R. T. (2005). "SUMO: A history of modification." Molecular Cell 18(1): 1-12.

Hay, R. T. (2006). "Role of ubiquitin-like proteins in transcriptional regulation." Ernst Schering Research Foundation workshop(57): 173-192.

Hay, R. T. (2007). "SUMO-specific proteases: a twist in the tail." Trends in Cell Biology 17(8): 370-376.

Hay, R. T. (2013). "Decoding the SUMO signal." Biochemical Society Transactions 41(2): 463-473.

Huang M, Miao ZH, Zhu H, Cai YJ, Lu W, and Ding J. (2008). Chk1 and Chk2 are differentially involved in homologous recombination repair and cell cycle arrest in response to DNA double-strand breaks induced by camptothecins. *Mol Cancer Ther* **7:** 1440-1449

Hecker, C. M., M. Rabiller, et al. (2006). "Specification of SUMO1- and SUMO2-interacting motifs." Journal of Biological Chemistry 281(23): 16117-16127.

Helchowski, C. M., L. F. Skow, et al. (2013). "A small ubiquitin binding domain inhibits ubiquitin-dependent protein recruitment to DNA repair foci." Cell Cycle 12(24): 3749-3758.

Hendriks, I. A., R. C. D'Souza, et al. (2015). "System-wide identification of wild-type SUMO-2 conjugation sites." Nature Communications 6.

Hershko, A., E. Leshinsky, et al. (1984). "ATP-dependent degradation of ubiquitin-protein conjugates." Proceedings of the National Academy of Sciences of the United States of America 81(6 I): 1619-1623.

Hochstrasser, M. (2009). "Origin and function of ubiquitin-like proteins." Nature 458(7237): 422-429.

Hofmann, K. and L. Falquet (2001). "A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems." Trends in Biochemical Sciences 26(6): 347-350.

Hofmann, R. M. and C. M. Pickart (1999). "Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair." Cell 96(5): 645-653.

Hu, X., A. Paul, et al. (2012). "Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates." Journal of Biological Chemistry 287(30): 25510-25519.

Hu, Y., R. Scully, et al. (2011). "RAP80-directed tuning of BRCA1 homologous recombination function at ionizing radiation-induced nuclear foci." Genes and Development 25(7): 685-700.

Huang, L., E. Kinnucan, et al. (1999). "Structure of an E6AP-UbcH7 complex: Insights into ubiquitination by the E2-E3 enzyme cascade." Science 286(5443): 1321-1326.

Hutten, S., G. Chachami, et al. (2014). "A role for the cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase." Journal of Cell Science 127(5): 1065-1078.

Itahana, Y., E. T. H. Yeh, et al. (2006). "Nucleocytoplasmic shuttling modulates activity and ubiquitination-dependent turnover of SUMO-specific protease 2." Molecular and Cellular Biology 26(12): 4675-4689.

Iwai, K. and F. Tokunaga (2009). "Linear polyubiquitination: A new regulator of NF-ŒʃB activation." EMBO Reports 10(7): 706-713.

Jackson, S. P. and J. Bartek (2009). "The DNA-damage response in human biology and disease." Nature 461(7267): 1071-1078.

Jaffray, E. G. and R. T. Hay (2006). "Detection of modification by ubiquitin-like proteins." Methods 38(1): 35-38.

Jin, J., X. Li, et al. (2007). "Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging." Nature 447(7148): 1135-1138.

Kerscher, O., R. Felberbaum, et al. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. Annual Review of Cell and Developmental Biology. 22: 159-180.

Knobel, P. A., R. Belotserkovskaya, et al. (2014). "USP28 is recruited to sites of DNA damage by the tandem BRCT domains of 53BP1 but plays a minor role in double-strand break metabolism." Molecular and Cellular Biology 34(11): 2062-2074.

Komander, D., M. J. Clague, et al. (2009). "Breaking the chains: Structure and function of the deubiquitinases." Nature Reviews Molecular Cell Biology 10(8): 550-

563.

Kulathu, Y. and D. Komander (2012). "Atypical ubiquitylation-the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages." Nature Reviews Molecular Cell Biology 13(8): 508-523.

Lallemand-Breitenbach, V., M. Jeanne, et al. (2008). "Arsenic degrades PML or PML-RARŒ± through a SUMO-triggered RNF4/ ubiquitin-mediated pathway." Nature Cell Biology 10(5): 547-555.

Lallemand-Breitenbach, V., J. Zhu, et al. (2001). "Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor Œ± degradation." Journal of Experimental Medicine 193(12): 1361-1371.

Lee, S., Y. C. Tsai, et al. (2006). "Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5." Nature Structural and Molecular Biology 13(3): 264-271.

Li, M. L. and R. A. Greenberg (2012). "Links between genome integrity and BRCA1 tumor suppression." Trends in Biochemical Sciences 37(10): 418-424.

Liew, C. W., H. Sun, et al. (2010). "RING domain dimerization is essential for RNF4 function." Biochemical Journal 431(1): 23-29.

Lukas, J., C. Lukas, et al. (2011). "More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance." Nature Cell Biology 13(10): 1161-1169.

Lundin, C., K. Erixon, et al. (2002). "Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells." Molecular and Cellular Biology 22(16): 5869-5878.

Luo, K., H. Zhang, et al. (2012). "Sumoylation of MDC1 is important for proper DNA

damage response." EMBO Journal 31(13): 3008-3019.

Mahajan, R., C. Delphin, et al. (1997). "A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2." Cell 88(1): 97-107.

Maser, R. S., K. J. Monsen, et al. (1997). "hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks." Molecular and Cellular Biology 17(10): 6087-6096.

Matic, I. and R. T. Hay (2012). Detection and quantitation of SUMO chains by mass spectrometry. Methods in Molecular Biology. 832: 239-247.

Matsumoto, M. L., K. E. Wickliffe, et al. (2010). "K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody." Molecular Cell 39(3): 477-484.

Matunis, M. J., E. Coutavas, et al. (1996). "A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex." Journal of Cell Biology 135(6): 1457-1470.

Mauro, M., Rego, M. A., Boisvert, R. A., Esashi, F., Cavallo, F., Jasin, M., & Howlett, N. G. (2012). p21 promotes error-free replication-coupled DNA double-strand break repair. *Nucleic Acids Research*, *40*(17), 8348–8360.

McKenna, S., T. Moraes, et al. (2003). "An NMR-based model of the ubiquitin-bound human ubiquitin conjugation complex Mms2¬∑Ubc13: The structural basis for lysine 63 chain catalysis." Journal of Biological Chemistry 278(15): 13151-13158.

Metzger, M. B., V. A. Hristova, et al. (2012). "HECT and RING finger families of E3 ubiquitin ligases at a glance." Journal of Cell Science 125(3): 531-537.

Meulmeester, E., M. Kunze, et al. (2008). "Mechanism and Consequences for Paralog-Specific Sumoylation of Ubiquitin-Specific Protease 25." Molecular Cell

30(5): 610-619.

Mimura, Y., K. Takahashi, et al. (2010). "Two-step colocalization of MORC3 with PML nuclear bodies." Journal of Cell Science 123(12): 2014-2024.

Morris, J. R., C. Boutell, et al. (2009). "The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress." Nature 462(7275): 886-890.

Moudry, P., C. Lukas, et al. (2012). "Ubiquitin-activating enzyme UBA1 is required for cellular response to DNA damage." Cell Cycle 11(8): 1573-1582.

Mukhopadhyay, D., F. Ayaydin, et al. (2006). "SUSP1 antagonizes formation of highly SUMO2/3-conjugated species." Journal of Cell Biology 174(7): 939-949.

Nacerddine, K., F. Lehembre, et al. (2005). "The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice." Developmental Cell 9(6): 769-779.

Nijman, S. M. B., M. P. A. Luna-Vargas, et al. (2005). "A genomic and functional inventory of deubiquitinating enzymes." Cell 123(5): 773-786.

Niu, J., Y. Shi, et al. (2011). "LUBAC regulates NF-ŒʃB activation upon genotoxic stress by promoting linear ubiquitination of NEMO." EMBO Journal 30(18): 3741-3753.

Noon, A. T. and A. A. Goodarzi (2011). "53BP1-mediated DNA double strand break repair: Insert bad pun here." DNA Repair 10(10): 1071-1076.

Ogawa, H., T. Komatsu, et al. (2009). "Transcriptional suppression by transient recruitment of ARIP4 to sumoylated nuclear receptor Ad4BP/SF-1." Molecular Biology of the Cell 20(19): 4235-4245.

Ono, Y., S. I. Iemura, et al. (2013). "PLEIAD/SIMC1/C5orf25, a novel autolysis regulator for a skeletal-muscle-specific calpain, CAPN3, scaffolds a CAPN3

substrate, CTBP1." Journal of Molecular Biology 425(16): 2955-2972.

Owerbach, D., E. M. McKay, et al. (2005). "A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation." Biochemical and Biophysical Research Communications 337(2): 517-520.

Panier, S., Y. Ichijima, et al. (2012). "Tandem Protein Interaction Modules Organize the Ubiquitin-Dependent Response to DNA Double-Strand Breaks." Molecular Cell 47(3): 383-395.

Pichler, A., P. Knipscheer, et al. (2004). "The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type." Nature Structural and Molecular Biology 11(10): 984-991.

Pinato, S., C. Scandiuzzi, et al. (2009). "RNF168, a new RING finger, MIU-containing protein that modifies chromatin by ubiquitination of histones H2A and H2AX." BMC Molecular Biology 10.

Plechanovov, A., E. G. Jaffray, et al. (2012). "Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis." Nature 489(7414): 115-120.

Plechanovov, A., E. G. Jaffray, et al. (2011). "Mechanism of ubiquitylation by dimeric RING ligase RNF4." Nature Structural and Molecular Biology 18(9): 1052-1059.

Popov, N., M. Wanzel, et al. (2007). "The ubiquitin-specific protease USP28 is required for MYC stability." Nature Cell Biology 9(7): 765-774.

Poulsen, S. L., R. K. Hansen, et al. (2013). "RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response." The Journal of cell biology 201(6): 797-807.

Redman, K. L. and M. Rechtsteiner (1989). "Identification of the long ubiquitin extension as ribosomal protein S27a." Nature 338(6214): 438-440.

Renatus, M., S. G. Parrado, et al. (2006). "Structural Basis of Ubiquitin Recognition

by the Deubiquitinating Protease USP2." Structure 14(8): 1293-1302.

Reverter, D. and C. D. Lima (2005). "Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex." Nature 435(7042): 687-692.

Rodriguez, M. S., C. Dargemont, et al. (2001). "SUMO-1 Conjugation in Vivo Requires Both a Consensus Modification Motif and Nuclear Targeting." Journal of Biological Chemistry 276(16): 12654-12659.

Rodriguez, M. S., J. M. P. Desterro, et al. (1999). "SUMO-1 modification activates the transcriptional response of p53." EMBO Journal 18(22): 6455-6461.

Rojas-Fernandez, A., A. Plechanovov√°, et al. (2014). "SUMO chain-induced dimerization activates RNF4." Molecular Cell 53(6): 880-892.

Sacco, J. J., J. M. Coulson, et al. (2010). "Emerging roles of deubiquitinases in cancer-associated pathways." IUBMB Life 62(2): 140-157.

Saitoh, H. and J. Hinchey (2000). "Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3." Journal of Biological Chemistry 275(9): 6252-6258.

Schnellhardt, M., K. Uzunova, et al. (2012). Analysis of cellular SUMO and SUMO-ubiquitin hybrid conjugates. Methods in Molecular Biology. 832: 81-92.

Schulz, S., G. Chachami, et al. (2012). "Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions." EMBO Reports 13(10): 930-938.

Seeler, J. S., A. Marchio, et al. (2001). "Common properties of nuclear body protein SP100 and TIF1Œ± chromatin factor: Role of SUMO modification." Molecular and Cellular Biology 21(10): 3314-3324.

Shao, G., D. R. Lilli, et al. (2009). "The Rap80-BRCC36 de-ubiquitinating enzyme

complex antagonizes RNF8-Ubc13-dependent ubiquitination events at DNA double strand breaks." Proceedings of the National Academy of Sciences of the United States of America 106(9): 3166-3171.

Shao, G., J. Patterson-Fortin, et al. (2009). "MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks." Genes and Development 23(6): 740-754.

Shen, L., M. H. Tatham, et al. (2006). "SUMO protease SENP1 induces isomerization of the scissile peptide bond." Nature Structural and Molecular Biology 13(12): 1069-1077.

Shen, L. N., C. Dong, et al. (2006). "The structure of SENP1-SUMO-2 complex suggests a structural basis for discrimination between SUMO paralogues during processing." Biochemical Journal 397(2): 279-288.

Shen, L. N., M. C. Geoffroy, et al. (2009). "Characterization of SENP7, a SUMO-2/3-specific isopeptidase." Biochemical Journal 421(2): 223-230.

Shen, T. H., H. K. Lin, et al. (2006). "The Mechanisms of PML-Nuclear Body Formation." Molecular Cell 24(3): 331-339.

Shin, E. J., H. M. Shin, et al. (2012). "DeSUMOylating isopeptidase: A second class of SUMO protease." EMBO Reports 13(4): 339-346.

Sobhian, B., G. Shao, et al. (2007). "RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites." Science 316(5828): 1198-1202.

Song, J., Z. Zhang, et al. (2005). "Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: A reversal of the bound orientation." Journal of Biological Chemistry 280(48): 40122-40129.

Stevenson, L. F., A. Sparks, et al. (2007). "The deubiquitinating enzyme USP2a regulates the p53 pathway by targeting Mdm2." EMBO Journal 26(4): 976-986.

Stewart, G. S., S. Panier, et al. (2009). "The RIDDLE Syndrome Protein Mediates a Ubiquitin-Dependent Signaling Cascade at Sites of DNA Damage." Cell 136(3): 420-434.

Stucki, M., J. A. Clapperton, et al. (2005). "MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks." Cell 123(7): 1213-1226.

Suh, H. Y., J. H. Kim, et al. (2012). "Crystal structure of DeSI-1, a novel deSUMOylase belonging to a putative isopeptidase superfamily." Proteins: Structure, Function and Bioinformatics 80(8): 2099-2104.

Sun, H. and T. Hunter (2012). "Poly-small ubiquitin-like modifier (PolySUMO)-binding proteins identified through a string search." Journal of Biological Chemistry 287(50): 42071-42083.

Sun, H., Y. Liu, et al. (2014). "Multiple Arkadia/RNF111 structures coordinate its polycomb body association and transcriptional control." Molecular and Cellular Biology 34(16): 2981-2995.

Takahashi, K., N. Yoshida, et al. (2007). "Dynamic regulation of p53 subnuclear localization and senescence by MORC3." Molecular Biology of the Cell 18(5): 1701-1709.

Tammsalu, T., I. Matic, et al. (2014). "Proteome-wide identification of SUMO2 modification sites." Science Signaling 7(323).

Tatham, M. H., M. C. Geoffroy, et al. (2008). "RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation." Nature Cell Biology 10(5): 538-546.

Tatham, M. H. and R. T. Hay (2003). "Ubiquitin and Ubiquitin-like Modifiers: Conserved Mechanisms and Diverse Functions." Chemtracts 16(13): 759-782.

Tatham, M. H. and R. T. Hay (2009). FRET-based in vitro assays for the analysis of SUMO protease activities. Methods in Molecular Biology. 497: 253-268.

Tatham, M. H., E. Jaffray, et al. (2001). "Polymeric Chains of SUMO-2 and SUMO-3 are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9." Journal of Biological Chemistry 276(38): 35368-35374.

Tatham, M. H., S. Kim, et al. (2005). "Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection." Nature Structural and Molecular Biology 12(1): 67-74.

Tatham, M. H., S. Kim, et al. (2003). "Role of an N-terminal site of Ubc9 in SUMO-1, -2, and -3 binding and conjugation." Biochemistry 42(33): 9959-9969.

Tatham, M. H., I. Matic, et al. (2011). "Comparative proteomic analysis identifies a role for SUMO in protein quality control." Science Signaling 4(178).

Tatham, M. H., A. Plechanovov, et al. (2013). "Ube2W conjugates ubiquitin to Œ±-amino groups of protein N-termini." Biochemical Journal 453(1): 137-145.

Tatham, M. H., M. S. Rodriguez, et al. (2009). "Detection of protein SUMOylation in vivo." Nature Protocols 4(9): 1363-1371.

Tenno T, Fujiwara K, Tochio H, Iwai K, Morita EH, et al. (2004). “Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains.” Genes Cells 9:865–75

Thrower, J. S., L. Hoffman, et al. (2000). "Recognition of the polyubiquitin proteolytic signal." EMBO Journal 19(1): 94-102.

Tokunaga, F., S. I. Sakata, et al. (2009). "Involvement of linear polyubiquitylation of NEMO in NF-ŒʃB activation." Nature Cell Biology 11(2): 123-132.

Torres-Rosell, J., I. Sunjevaric, et al. (2007). "The Smc5-Smc6 complex and SUMO

modification of Rad52 regulates recombinational repair at the ribosomal gene locus." Nature Cell Biology 9(8): 923-931.

Uzunova, K., K. Gottsche, et al. (2007). "Ubiquitin-dependent proteolytic control of SUMO conjugates." Journal of Biological Chemistry 282(47): 34167-34175.

Valero, R., M. Bayes, et al. (2001). "Characterization of alternatively spliced products and tissue-specific isoforms of USP28 and USP25." Genome biology 2(10).

Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C, Fushman D. (2004). “Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signalling”. J. Biol. Chem. 279:7055–63

Vertegaal, A. C. O., J. S. Andersen, et al. (2006). "Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics." Molecular and Cellular Proteomics 5(12): 2298-2310.

Wang, B. and S. J. Elledge (2007). "Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage." Proceedings of the National Academy of Sciences of the United States of America 104(52): 20759-20763.

Wang, B., S. Matsuoka, et al. (2007). "Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response." Science 316(5828): 1194-1198.

Wang, L., C. Wansleeben, et al. (2014). "SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development." EMBO Reports 15(8): 878-885.

Wasylishen, A. R., M. Chan-Seng-Yue, et al. (2013). "MYC phosphorylation at novel regulatory regions suppresses transforming activity." Cancer Research 73(21): 6504-6515.

Weger, S., E. Hammer, et al. (2005). "Topors acts as a SUMO-1 E3 ligase for p53 in vitro and in vivo." FEBS Letters 579(22): 5007-5012.

Wickliffe, K. E., S. Lorenz, et al. (2011). "The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2." Cell 144(5): 769-781.

Wong, K. A., R. Kim, et al. (2004). "Protein inhibitor of activated STAT Y (PIASy) and a splice variant lacking exon 6 enhance sumoylation but are not essential for embryogenesis and adult life." Molecular and Cellular Biology 24(12): 5577-5586.

Wu, J., C. Liu, et al. (2012). "RAP80 protein is important for genomic stability and is required for stabilizing BRCA1-A complex at DNA damage sites in vivo." Journal of Biological Chemistry 287(27): 22919-22926.

Wu, Y., Y. Wang, et al. (2013). "The Deubiquitinase USP28 Stabilizes LSD1 and Confers Stem-Cell-like Traits to Breast Cancer Cells." Cell Reports 5(1): 224-236.

Wu-Baer, F., K. Lagrazon, et al. (2003). "The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin." Journal of Biological Chemistry 278(37): 34743-34746.

Xu, P., D. M. Duong, et al. (2009). "Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation." Cell 137(1): 133-145.

Xu, Y., A. Plechanovov, et al. (2014). "Structural insight into SUMO chain recognition and manipulation by the ubiquitin ligase RNF4." Nature Communications 5.

Xu, Z. and S. W. N. Au (2005). "Mapping residues of SUMO precursors essential in differential maturation by SUMO-specific protease, SENP1." Biochemical Journal 386(2): 325-330.

Xu, Z. and S. W. N. Au (2005). "Mapping residues of SUMO precursors essential in differential maturation by SUMO-specific protease, SENP1." Biochemical Journal 386(2): 325-330.

Yan, J. and A. M. Jetten (2008). "RAP80 and RNF8, key players in the recruitment of

repair proteins to DNA damage sites." Cancer Letters 271(2): 179-190.

Yan, J., X. P. Yang, et al. (2007). "RAP80 interacts with the SUMO-conjugating enzyme UBC9 and is a novel target for sumoylation." Biochemical and Biophysical Research Communications 362(1): 132-138.

Yan, Z., Y. S. Kim, et al. (2002). "RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor." Journal of Biological Chemistry 277(35): 32379-32388.

Yang, L., J. R. Mullen, et al. (2006). "Purification of the yeast Slx5-Slx8 protein complex and characterization of its DNA-binding activity." Nucleic Acids Research 34(19): 5541-5551.

Yin, Y., A. Seifert, et al. (2012). "SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage." Genes and Development 26(11): 1196-1208.

Young, P., Q. Deveraux, et al. (1998). "Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a." Journal of Biological Chemistry 273(10): 5461-5467.

Yu, X., L. C. Wu, et al. (1998). "The C-terminal (BRCT) domains of BRCA1 interact in vive with CtIP, a protein implicated in the CtBP pathway of transcriptional repression." Journal of Biological Chemistry 273(39): 25388-25392.

Yun, M. H. and K. Hiom (2009). "Understanding the functions of BRCA1 in the DNA-damage response." Biochemical Society Transactions 37(3): 597-604.

Zhang, D., K. Zaugg, et al. (2006). "A Role for the Deubiquitinating Enzyme USP28 in Control of the DNA-Damage Response." Cell 126(3): 529-542.

Zhang, F. P., L. Mikkonen, et al. (2008). "Sumo-1 function is dispensable in normal mouse development." Molecular and Cellular Biology 28(17): 5381-5390.

Zhen, Y., P. A. Knobel, et al. (2014). "Regulation of USP28 deubiquitinating activity by SUMO conjugation." Journal of Biological Chemistry 289(50): 34838-34850.